Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US05/006505

International filing date: 23 February 2005 (23.02.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/547,256
Filing date: 23 February 2004 (23.02.2004)

Date of receipt at the International Bureau: 07 April 2005 (07.04.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

1302262

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

March 31, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/547,256*
FILING DATE: *February 23, 2004*
RELATED PCT APPLICATION NUMBER: *PCT/US05/06505*

Certified by

Jon W. Dudas

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

Please type a plus sign (+) inside this box → [+]

PTO/SB/16 (5-02)
Approved for use through 4/30/2003. OMB 0651-0032.
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

INVENTOR(S)

Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Helge	Zieler	Chicago, IL
Rongguan	Jin	Chicago, IL
Jennifer M.	Mach	Chicago, IL
Gary	Rudgers	Chicago, IL

☒ *Additional inventors are being named on the* 1 *separately numbered sheets attached hereto*

TITLE OF THE INVENTION (280 characters max)

PLANTS TRANSFORMED WITH MINI-CHROMOSOMES

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☒ Customer Number **24573** →

OR *Type Customer Number here*

Place Customer Number Bar Code Label here 24573

PATENT TRADEMARK OFFICE

☐ Firm *or* Individual Name					
Address					
Address					
City		State		ZIP	
Country		Telephone		Fax	

ENCLOSED APPLICATION PARTS ***(check all that apply)***

☒ Specification *Number of Pages* **31**

☒ Drawing(s) *Number of Sheets* **12**

☐ Application Data Sheet. See 37 CFR 1.76

☐ CD(s), Number

☐ Other (specify)

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT *(check one)*

☒ Applicant claims small entity status. See 37 CFR 1.27.

☐ A check or money order is enclosed to cover the filing fees

☒ The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number **02-1818**

☐ Payment by credit card. Form PTO-2038 is attached.

FILING FEE AMOUNT ($) **$80.00**

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ No.

☐ Yes, the name of the U.S. Government agency and the Government contract number are: _____

Respectfully submitted,

SIGNATURE Tin-Chuen Yeung

Date **2/23/2004**

TYPED or PRINTED NAME **Tin-Chuen Yeung**

REGISTRATION NO. *(if appropriate)* **40,240**

Docket Number: **112283-060**

TELEPHONE **(312) 807-4369**

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is used by the public to file (and by the PTO to process) a provisional application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the complete provisional application to the PTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Mail Stop Provisional Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need asssitance in completing the form, call 1-800-PTO-9199 and select option 2.

PROVISIONAL APPLICATION COVER SHEET
Additional Page

PTO/SB/16 (8-00)
Approved for use through 10/31/2002. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Docket Number	112283-060	Type a plus sign (+) inside this box →	+

INVENTOR(S)/APPLICANT(S)

Given Name (first and middle [if any])	Family or Surname	Residence (City and either State or Foreign Country)
Daphne	**Preuss**	**Chicago, IL**
Mich B.	**Hein**	**Chicago, IL**
Greg P.	**Copenhaver**	**Chapel Hill, NC**
Kevin C.	**Keith**	**Chicago, IL**

Number 1 of 1

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s): Helge Zieler et al.
Appl. No.: Not yet assigned
Filed: February 23, 2004
Title: PLANTS TRANSFORMED WITH MINI-CHROMOSOMES
Art Unit: Not yet assigned
Examiner: Not yet assigned
Docket No.: 112283-060

Mail Stop PROVISIONAL APPLICATION
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

CERTIFICATE OF MAILING BY EXPRESS MAIL UNDER 37 CFR 1.10

Sir:

I hereby certify that the following documents relating to the above-identified application:

1. Provisional Application for Patent Cover Sheet (2 pgs – in duplicate);
2. Specification (31 pages);
3. Drawings (7 Figures, 12 sheets); and
4. Return Receipt Postcard.

are being deposited with the United States Postal Service with sufficient postage as Express Mail in an envelope addressed to:

Mail Stop PROVISIONAL APPLICATION
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

on February 23, 2004.

Respectfully submitted,

BELL, BOYD & LLOYD LLC

Judith L. Goldberg
Name of Person Mailing Correspondence


Signature

EV351200667US
Express Mail Mailing Label Number

SPECIFICATION

TITLE OF THE INVENTION

"Plants transformed with mini-chromosomes"

BACKGROUND OF THE INVENTION

Two general approaches are used for introduction of new genetic information ("transformation") into cells. One approach is to introduce the new genetic information as part of another DNA molecule, referred to as an "episomal vector," or "mini-chromosome", which can be maintained as an independent unit (an episome) apart from the host chromosomal DNA molecule(s). Episomal vectors contain all the necessary DNA sequence elements required for DNA replication and maintenance of the vector within the cell. Many episomal vectors are available for use in bacterial cells (for example, see Maniatis *et al.*, 1982). However, only a few episomal vectors that function in higher eukaryotic cells have been developed. Higher eukaryotic episomal vectors were primarily based on naturally occurring viruses. In higher plant systems gemini viruses are double-stranded DNA viruses that replicate through a double-stranded intermediate upon which an episomal vector could be based, although the gemini virus is limited to an approximately 800 bp insert. Although an episomal plant vector based on the Cauliflower Mosaic Virus has been developed, its capacity to carry new genetic information also is limited (Brisson *et al.*, 1984).

The other general method of genetic transformation involves integration of introduced DNA sequences into the recipient cell's chromosomes, permitting the new information to be replicated and partitioned to the cell's progeny as a part of the natural chromosomes. The introduced DNA usually is broken and joined together in various combinations before it is integrated at random sites into the cell's chromosome (see, for example Wigler *et al.*, 1977). Common problems with this procedure are the rearrangement of introduced DNA sequences and unpredictable levels of expression due to the location of the transgene in the genome or so called "position effect variation" (Shingo *et al.*, 1986). Further, unlike episomal DNA, integrated DNA cannot normally be precisely removed. A more refined form of integrative transformation can be achieved by

exploiting naturally occurring viruses that integrate into the host's chromosomes as part of their life cycle, such as retroviruses (see Cepko *et al.*, 1984). In mouse, homologous integration has recently become common, although it is significantly more difficult to use in plants (Lam *et al.* 1996).

The most common genetic transformation method used in higher plants is based on the transfer of bacterial DNA into plant chromosomes that occurs during infection by the phytopathogenic soil bacterium *Agrobacterium* (see Nester *et al.*, 1984). By substituting genes of interest for the naturally transferred bacterial sequences (called T-DNA), investigators have been able to introduce new DNA into plant cells. However, even this more "refined" integrative transformation system is limited in three major ways. First, DNA sequences introduced into plant cells using the *Agrobacterium* T-DNA system are frequently rearranged (see Jones *et al.*, 1987). Second, the expression of the introduced DNA sequences varies between individual transformants (see Jones *et al.*, 1985). This variability is presumably caused by rearranged sequences and the influence of surrounding sequences in the plant chromosome (*i.e.*, position effects), as well as methylation of the transgene. A third drawback of the *Agrobacterium T-DNA* system is the reliance on a "gene addition" mechanism: the new genetic information is added to the genome (*i.e.*, all the genetic information a cell possesses) but does not replace information already present in the genome.

One attractive alternative to commonly used methods of transformation is the use of an artificial chromosome. Artificial chromosomes are man-made linear or circular DNA molecules constructed from cis-acting DNA sequence elements that provide replication and partitioning of the constructed chromosomes (see Murray *et al.*, 1983). Desired elements include: (1) Autonomous Replication Sequences (ARS) (these have properties of replication origins, which are the sites for initiation of DNA replication), (2) Centromeres (site of kinetochore assembly and responsible for proper distribution of replicated chromosomes at mitosis or meiosis), and (3) if the chromosome is linear, telomeres (specialized DNA structures at the ends of linear chromosomes that function to stabilize the ends and facilitate the complete replication of the extreme termini of the DNA molecule).

The essential chromosomal elements for construction of artificial chromosomes have been precisely characterized in lower eukaryotic species, and more recently in mouse and human. ARSs have been isolated from unicellular fungi, including *Saccharomyces cerevisiae* (brewer's yeast) and *Schizosaccharomyces pombe* (see Stinchcomb *et al.*, 1979 and Hsiao *et al.*, 1979). An ARS behaves like a replication origin allowing DNA molecules that contain the ARS to be replicated as an episome after introduction into the cell nuclei of these fungi. DNA molecules containing these sequences replicate, but in the absence of a centromere they are partitioned randomly into daughter cells.

Artificial chromosomes have been constructed in yeast using the three cloned essential chromosomal elements. Murray *et al.*, 1983, disclose a cloning system based on the in vitro construction of linear DNA molecules that can be transformed into yeast, where they are maintained as artificial chromosomes. These yeast artificial chromosomes (YACs) contain cloned genes, origins of replication, centromeres and telomeres and are segregated in daughter cells with high fidelity when the YAC is at least 100 kB in length. Smaller CEN-containing vectors may be stably segregated, however, when in circular form.

None of the essential components identified in unicellular organisms, however, function in higher eukaryotic systems. For example, a yeast CEN sequence will not confer stable inheritance upon vectors transformed into higher eukaryotes. While such DNA fragments can be readily introduced, they do not stably exist as episomes in the host cell. This has seriously hampered efforts to produce artificial chromosomes in higher organisms.

In one case, a plant artificial chromosome was discussed (Richards *et al.*, U.S. Patent No. 5,270,201). However, this vector was based on plant telomeres, as a functional plant centromere was not disclosed. While telomeres are important in maintaining the stability of chromosomal termini, they do not encode the information needed to ensure stable inheritance of an artificial chromosome. It is well documented that centromere function is crucial for stable chromosomal inheritance in almost all eukaryotic organisms (reviewed in Nicklas 1988). For example, broken chromosomes that lack a centromere (acentric chromosomes) are rapidly lost from cell lines, while fragments that have a

centromere are faithfully segregated. The centromere accomplishes this by attaching, via centromere binding proteins, to the spindle fibers during mitosis and meiosis, thus ensuring proper gene segregation during cell divisions.

In contrast to the detailed studies done in *S. cerevisiae* and *S. pombe*, less is known about the molecular structure of functional centromeric DNA of higher eukaryotes. Ultrastructural studies indicate that higher eukaryotic kinetochores, which are specialized complexes of proteins that form on the centromere during late prophase, are large structures (mammalian kinetochore plates are approximately 0.3 μm in diameter) which possess multiple microtubule attachment sites (reviewed in Rieder, 1982). It is therefore possible that the centromeric DNA regions of these organisms will be correspondingly large, although the minimal amount of DNA necessary for centromere function may be much smaller.

The above studies have been useful in elucidating the structure and function of centromeres-. The extensive literature indicating both the necessity of centromeres for stable inheritance of chromosomes, and the non-functionality of yeast centromeres in higher organisms, demonstrate that cloning of a functional centromere from a higher eukaryote is a necessary first step in the production of artificial chromosomes suitable for use in higher plants and animals. The production of artificial chromosomes with centromeres which function in higher eukaryotes would overcome many of the problems associated with the prior art and represent a significant breakthrough in biotechnology research.

These and other aspects and attributes of the present invention will be discussed with reference to the following drawings and accompanying specification.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a list of sequences relevant to *Brassica* mini-chromosome BR5R4-1;

Figure 2 is a list of examples of non-plant derived promoters for transforming plant cells;

Figure 3 is an example of a mini-chromosome vector in the present invention containing 2 genes;

Figure 4 is another example of a mini-chromosome vector in the present invention containing 4 genes;

Figure 5 is a mini-chromosome from which all bacterial sequences have been removed. In this embodiment, bacterial sequence present between or among the plant-expressed genes or other mini-chromosome sequences would be excised prior to removal of the remaining bacterial sequences, by cutting the mini-chromosome with homing endonuclease #1, and re-ligating the structure such that the antibiotic-resistance gene #1 has been lost;

Figure 6 shows various structural configurations by which mini-chromosome elements can be oriented with respect to each other; and

Figure 7 shows pictures of a mini-chromosome transformed plant and parts thereof.

DETAILED DESCRIPTION OF THE INVENTION

While this invention is susceptible of embodiment in many different forms, there is shown in the drawing, and will be described herein in detail, specific embodiments thereof with the understanding that the present disclosure is to be considered as an exemplification of the principles of the invention and is not intended to limit the invention to the specific embodiments illustrated.

The invention is generally related to methods of generating plants transformed with novel autonomous mini-chromosomes. Mini-chromosomes with novel compositions and structures are used to transform plants cells which are in turn used to generate the plant. Methods for generating the plant include methods for delivering the mini-chromosome into plant cell to transform the cell, methods for selecting the transformed cell, and methods for isolating plants transformed with the mini-chromosome. Plants generated in the present invention contain novel genes introduced into their genome by integration into existing chromosomes.

One aspect of the present invention relates to centromere compositions in the mini-chromosome. Another aspect of the present invention relates to the novel structure of the mini-chromosome, particularly structures lacking bacterial sequences. Yet another aspect

of the present invention is related to methods for generating variants of centromeres by passage through bacterial or plant hosts. Still another aspect of the present invention is related to methods of making and composition of non-plant promoters for expressing genes in plants. A further aspect is related to methods for isolating plants transformed with mini-chromosomes (including regeneration method, DNA delivery, selection of transformed cells). Still yet another aspect is related to methods for using exonuclease to enrich for circular mini-chromosome DNA in genomic DNA preparations Still another further aspect is related to methods for assessing mini-chromosome performance (lineage-baased inheritance assays, use of chromosome loss agents to demonstrate autonomy)

The advantages of the present invention include: autonomous, independent linkage group for accelerating breeding; lack of disruption of host genome; unlimited payload for genes; defined genetic context for predictable gene expression; higher transformation frequency due to elimination of inefficient integration step; and the ability to eliminate mini-chromosomes in any tissues.

I. Composition of mini-chromosomes and mini-chromosome construction

The mini-chromosomes of the present invention are constructed to include various components which are novel, which include, but are not limited to, the centromere comprising novel repeating centromeric sequences, and the promoters, particularly promoters derived from non-plant species. An example of the mini-chromosome in the present invention is the *Brassica* BB5R4-1 mini-chromosome. The sequences (SEQ ID NOS:1 to 3) relevant to the BB5R4-1 mini-chromosomes are listed in Figure 1.

Centromere composition

The centromere in the mini-chromosome of the present invention comprises novel repeating centromeric sequences. For example, the centromere of the BB5R4-1 mini-chromosome is 64kb of Brassica centromere DNA as determined by CHEF gel analysis. To determine the sequence composition of the centromere, we randomly sheared the mini-chromosome and cloned small fragments for sequencing, obtaining 11,010 bases of sequence from the centromere insert, a 0.17x coverage of the centromere. Of this sequence 9,533 bases were composed of centromere satellite repeat, the consensus of

which is shown in Figure 1b (SEQ ID NO:2). The satellite repeat was found to be 180+/-2 bp long. The remaining 1,477 bases of mini-chromosome sequence covered a unique sequence shown in Figure 1c (SEQ ID NO:3). This sequence is considered a sampling of the centromere content of BB5R4-1.

Mini-chromosome sequence content and structure

The mini-chromosome vector of the present invention may contain two or more plant-expressed genes, such as those shown in Figures 3 and 4. Each gene consists of a promoter, a coding region and a terminator sequence, which are separated from each other by restriction endonuclease sites or recombination sites or both. Genes may also include introns, which may be present in any number and at any position within the transcribed portion of the gene, incuding the 5' untranslated sequence, the coding region and the 3' untranslated sequence. Introns may be natural plant introns derived from any plant, or artificial introns based on the splice site consensus that has been defined for plant species. The coding regions of the genes can encode any protein, including but not limited to visible marker genes (fluorescent protein genes, other genes conferring a visible phenotype to the plant) or selectable marker genes (conferring resistance to antibiotics, herbicides or other toxic compounds or encoding a protein that confers a growth advantage to the cell expressing the protein) or genes which confer some commercial or agronomic value to the transformed plant. Multiple genes can be placed on the same mini-chromosome vector, limited only by the number of restriction endonuclease sites or site-specific recombination sites present in the vector. The genes are separated from each other by restriction endonuclease sites, homing endonucelase sites, recombination sites or any combinations thereof. Any number of genes can be present: Figures 3 and 4 show mini-chromosome vector structures with 2 and 4 genes, respectively.

The mini-chromosome vector also contains a bacterial plasmid backbone for propagation of the plasmid in bacteria such as E. coli. The plasmid backbone may be that of a low-copy vector or in other embodiments it may be desirable to use a mid to high level copy backbone. In one embodiment of the invention, this backbone contains the replicon of the F' plasmid of E. coli. However, other plasmid replicons, such as the bacteriophage P1 replicon, or other low-copy plasmid systems such as the RK2 replication

origin, may also be used. The backbone may include one or several antibiotic-resistance genes conferring resistance to a specific antibiotic to the bacterial cell in which the plasmid is present. Bacterial antibiotic-resistance genes include but are not limited to kanamycin-, ampicillin-, chloramphenicol-, streptomycin-, spectinomycin-, tetracycline- and gentamycin-resistance genes.

The mini-chromosome vector may also contain plant telomeres, as well as additional "stuffer DNA" sequences that serve to separate the various components on the mini-chromosome (centromere, genes, telomeres) from each other. The stuffer DNA may be of any origin, prokaryotic or eukaryotic, and from any genome or species, plant, animal, microbe or organelle, or may be of synthetic origin. The stuffer DNA can range from 100 bp to 10 Mb in length and can be repetitive in sequence, with unit repeats from 10 to 1,000,000 bp. Examples of repetitive sequences that can be used as stuffer DNAs include but are not limited to: rDNA, satellite repeats, retroelements, transposons, pseudogenes, transcribed genes, microsatellites, short sequence repeats and combinations thereof. Alternatively, the stuffer DNA can consist of unique, non-repetitive DNA of any origin or sequence. The stuffer sequences may also include DNA with the ability to form boundary domains, such as but not limited to scaffold attachment regions (SARs) or matrix attachment regions (MARs)

In one embodiment of the invention, the mini-chromosome has a circular structure without telomeres, as shown in Figures 3 and 4 "circular". In another embodiment, the mini-chromosome has a circular structure with telomeres, as shown in Figures 3 and 4 "linear". In a third embodiment, the mini-chromosome has a linear structure with telomeres, as would result if the "linear" structure shown in Figures 3 and 4 were to be cut with homing endonuclease, exposing the telomeres at the ends of a DNA molecule that contains all of the sequence contained in the original, closed construct with the exception of the antibiotic-resistance gene #1. In a fourth embodiment of the invention, the telomeres could be placed in such a manner that the bacterial replicon, backbone sequences, antibiotic-resistance genes and any other sequences of bacterial origin and present for the purposes of propagation of the mini-chromosome in bacteria, can be removed from the plant-expressed genes, the centromere, telomeres, and other sequences

by cutting the structure with homing endonuclease #2 (Figure 5). This results in a mini-chromosome from which all bacterial sequences have been removed. In this embodiment, bacterial sequence present between or among the plant-expressed genes or other mini-chromosome sequences would be excised prior to removal of the remaining bacterial sequences, by cutting the mini-chromosome with homing endonuclease #1, and re-ligating the structure such that the antibiotic-resistance gene #1 has been lost (Figure 5). In all of the structures shown in figures 3, 4 and 5, homing endonucleases and their sites can be replaced with any rare-cutting endonuclease or recombinase and its specific recognition site, as long as that site is only present at the indicated positions in the mini-chromosomes.

Various structural configurations are possible by which mini-chromosome elements can be oriented with respect to each other. A simple set of examples are shown in Figure 6. A centromere can be placed on a mini-chromosome either between genes (Fig 6A) or outside a cluster of genes next to one telomere (Fig. 6B) or next to the other telomere (Fig 6C). Stuffer DNAs can be combined with these configurations to place the stuffer sequences inside the telomeres (Figure 6D), around the centromere (Fig. 6E), between genes (Fig. 6F) or any combination thereof (not shown). Thus, a large number of alternative mini-chromosome structures are possible, depending on the relative placement of centromere DNA, genes, stuffer DNAs, bacterial sequences, telomeres, and other sequences. The sequence content of each of these variants is the same, but their structure is different depending on how the sequences are placed. These variations in architecture are possible both for linear and for circular mini-chromosomes, although only linear structures are diagrammed in Figure 6.

Use of promoter regions isolated from *Drosophila melanogaster* and *Saccharomyces cerevisiae* to express genes in plants

The promoter in the mini-chromosome of the present invention can be derived from plant or non-plant species. In a preferred embodiment, the nucleotide sequence of the promoter is derived from non-plant species for the expression of genes in either dicotyledon plant cells such as tobacco, tomato, potato, soybean, canola, sunflower, alfalfa, cotton and *Arabidopsis*, or monocotyledonous plant cell, such as wheat, maize, rye, rice, turf grass, oat, barley, sorghum, millet, and sugarcane. In one embodiment, the

non-plant promoters are derived from *Drosophila melanogaster* or *Saccharomyces cerevisiae*. Table 1 lists the promoters from *Drosophila melanogaster* and *Saccharomyces cerevisiae* are used to derive the examples of non-plant promoters in the present invention. Promoters derived from any animal, protest, or fungi are also contemplated. Figure 2 lists examples of promoter sequences derived from *Drosophila melanogaster* or *Saccharomyces cerevisiae* (SEQ ID NOS:4 to 23). These non-plant promoters can be operably linked to nucleic acid sequences encoding polypeptides or non-protein-expressing sequences including, but not limited to, antisense RNA and ribozymes, to form nucleic acid constructs, vectors, and host cells (prokaryotic or eukaryotic), comprising the promoters.

Table 1

Drosophila melanogaster Promoters
http://flybase.bio.indiana.edu/

Seq No.	Symbol	Flybase ID	Standard promoter gene name	Gene Product	Chromosome
4	Pgd	FBgn0004654	Phosphogluconate dehydrogenase	6-phosphogluconate dehydrogenase	X
5	Grim	FBgn0015946	grim	grim-P138	3
5	Uro	FBgn0003961	Urate oxidase	Uro-P1	2
7	Sna	FBgn0003448	snail	sna-P1	2
8	Rh3	FBgn0003249	Rhodopsin 3	Rh3	3
9	Lsp-1 γ	FBgn0002564	Larval serum protein 1 γ	Lsp1γ-P1	3

Saccharomyces cerevisiae Promoters
Information from:
http://www.yeastgenome.org/SearchContents.shtml

Seq No.	Symbol	Systematic Name	Standard promoter gene name	Gene Product	Chromosome
10	Tef-2	YBR118W	TEF2 (Translation elongation factor promtoer)	Translation elongation factor EF-1 alpha	2
11	Leu-1	YGL009C	LEU1 (LEUcine biosynthesis)	isopropylmalate isomerase	7
12	Met16	YPR167C	METhionine requiring	3'phosphoadenylyls ulfate reductase	16
13	Leu-2	YCL018W	LEU2 (leucine biosynthesis)	beta-IPM (isopropylmalate) dehydrogenase	3
14	His-4	YCL030C	HIS4 (HIStidine requiring)	histidinol dehydrogenase	3
15	Met-2	YNL277W	MET2 (methionine requiring)	L-homoserine-O-acetyltransferase	14
16	Ste-3	YKL178C	STE3 (alias DAF2 Sterile)	a-factor receptor	11

17	Arg-1	YOL058W	ARG1(alias ARG10 ARGinine requiring)	arginosuccinate synthetase	15
18	Pgk-1	YCR012W	PGK1 (phosphoglycerate kinase)	phosphoglycerate kinase	3
19	GPD-1	YDL022W	GPD1 (alias DAR1/HOR1/OSG1/ OSR5: glycerol-3-phosphate dehydrogenase activity	glycerol-3-phosphate dehydrogenase	4
20	ADH1	YOL086C	ADH1 (alias ADC1)	alcohol dehydrogenase	15
21	GPD-2	YOL059W	GPD2 (alias GPD3: glycerol-3-phosphate dehydrogenase activity	glycerol-3-phosphate dehydrogenase	15
22	Arg-4	YHR018C	ARGinine requiring	argininosuccinate lyase	8
23	Yat-1	YAR035W	YAT-1(carnitine acetyltransferase)	carnitine acetyltransferase	1

The present invention relates to methods for producing a polypeptide, comprising of cultivating plant material for the production of the polypeptide at any level, wherein the plant host cells comprises a first nucleic acid sequence encoding the polypeptide operably linked to a second nucleic acid sequence comprising a promoter foreign to the nucleic acid sequence, wherein the promoter comprises a sequence selected from the group consisting of SEQ ID NOS:4 to 23 (Figure 2) and subsequences thereof; and mutant, hybrid, and tandem promoters thereof.

The present invention also relates to methods for producing non-protein expressed sequences, comprising of cultivating plant material for the production of the non-protein expressed sequence, wherein the plant host cell comprises a first nucleic acid sequence encoding the non-protein expressed sequences operably linked to a second nucleic acid sequence comprising a promoter foreign to the nucleic acid sequence, wherein the promoter comprises a sequence selected from the group consisting of SEQ ID NOS:4 to 23 (Figure 2) and subsequences thereof; and mutant, hybrid, and tandem promoters thereof.

The present invention also relates to isolated promoter sequences and to constructs, vectors, and plant host cells comprising one or more of the promoters operably

linked to a nucleic acid sequence encoding a polypeptide or non-protein expressing sequence.

In the methods of the present invention, the promoter may also be a mutant of the promoters having a substitution, deletion, and/or insertion of one or more nucleotides in the nucleic acid sequence of SEQ ID NOS:4 to 23.

The present invention also relates to methods for obtaining derivative promoters of SEQ ID NOS:4 to 23.

The techniques used to isolate or clone a nucleic acid sequence comprising a promoter of interest are known in the art and include isolation from genomic DNA. The cloning procedures may involve excision or amplification, for example by polymerase chain reaction, and isolation of a desired nucleic acid fragment comprising the nucleic acid sequence encoding the promoter, insertion of the fragment into a vector molecule, and incorporation of the recombinant vector into the plant cell.

The term "promoter" is defined herein as a DNA sequence that allows the binding of RNA polymerase (including but not limited to RNA polymerase I, RNA polymerase II and RNA polymerase II from eukaryotes) and directs the polymerase to a downstream transcriptional start site of a nucleic acid sequence encoding a polypeptide to initiate transcription. RNA polymerase effectively catalyzes the assembly of messenger RNA complementary to the appropriate DNA strand of the coding region. The term "promoter" will also be understood to include the 5' non-coding region (between promoter and translation start) for translation after transcription into mRNA, cis-acting transcription control elements such as enhancers, and other nucleotide sequences capable of interacting with transcription factors.

The term "derivative promoter" is defined herein as a promoter having a nucleotide sequence comprising a substitution, deletion, insertion and/or rearrangement of one or more nucleotides of a parent promoter, wherein the mutant promoter has more or less promoter activity or a different specificity compared to the corresponding parent promoter. The term "derivative promoter" will also encompass natural variants and in vitro generated variants obtained using methods well known in the art such as classical mutagenesis, site-directed mutagenesis, and DNA shuffling.

The term "hybrid promoter" is defined herein as parts of two or more promoters that are fused together to generate a sequence that is a fusion of the two or more promoters, which is operably linked to a coding sequence and mediates the transcription of the coding sequence into mRNA.

The term "tandem promoter" is defined herein as two or more promoter sequences each of which is operably linked to a coding sequence and mediates the transcription of the coding sequence into mRNA.

The term "operably linked" is defined herein as a configuration in which a control sequence, e.g., a promoter sequence, is appropriately placed at a position relative to a coding sequence such that the control sequence directs the production of a polypeptide encoded by the coding sequence.

The term "coding sequence" is defined herein as a nucleic acid sequence that is transcribed into mRNA which is translated into a polypeptide when placed under the control of promoter sequences. The boundaries of the coding sequence are generally determined by the ATG start codon located just upstream of the open reading frame at the 5' end of the mRNA and in some cases, a transcription terminator sequence located just downstream of the open reading frame at the 3' end of the mRNA. A coding sequence can include, but is not limited to, genomic DNA, cDNA, semisynthetic, synthetic, and recombinant nucleic acid sequences.

The term "non-protein-expressing sequence" is defined herein as a nucleic acid sequence that is or is not transcribed into RNA were the transcribed nucleotides function within the host cells as, but not limited to, ribozymes or antisense RNA.

The term "polypeptide" does not refer to a specific length of the encoded product and, therefore, encompasses peptides, oligopeptides, and proteins. The term "heterologous polypeptide" is defined as a polypeptide which is not native to the plant cell, a native polypeptide in which modifications have been made to alter the native sequence, or a native polypeptide whose expression is quantitatively altered as a result of a manipulation of the plant cell by recombinant DNA techniques.

Constructing mini-chromosomes by site-specific recombination

This invention describes the construction of plant mini-chromosomes using site-specific recombination sequences (for example those recognized by the bacteriophage P1 Cre recombinase, or the bacteriophage lambda integrase, or similar recombination enzymes). A compatible recombination site, or a pair of such sites, is present on both the centromere containing DNA clones and the donor DNA clones. Incubation of the donor clone and the centromere clone in the presence of the recombinase enzyme causes strand exchange to occur between the recombination sites in the two plasmids; the resulting mini-chromosomes contain centromere sequences as well as mini-chromosome vector sequences. The DNA molecules formed in such recombination reactions is transformed into *E. coli*, other bacteria, yeast or plant cells by common methods in the field including, but not limited to, chemical transformation, electroporation, particle bombardment, whiskers, or other transformation method followed by selection for marker genes including chemical, enzymatic, color, or other marker present on either parental plasmid, allowing for the selection of transformants harboring mini-chromosomes.

Centromere variants arising from passage through bacteria, plant or other hosts or processes

In the methods of the present invention, the resulting mini-chromosome DNA sequence may also be a derivative of the donor clone or centromere clone having substitutions, deletions, insertions, duplications and/or rearrangments of one or more nucleotides in the nucleic acid sequence. Such nucleotide mutations may occur individually or consecutively in stretches of 1, 2, 3, 4, 5, 10, 20, 40, 80, 100, 200, 400, 800, 1000, 2000, 4000, 8000, 10000, 50000, 100000, and about 200000, including all ranges in-between.

Variations of mini-chromosomes may arise through passage of mini-chromosomes through various hosts including virus, bacteria, yeast, plant or other prokaryotic or eukaryotic organism and may occur through passage of multiple hosts or individual host. Variations may also occur by replicating the mini-chromosome in vitro.

Derivatives may be identified through sequence analysis, or variations in mini-chromosome molecular weight through electrophoresis such as, but not limited to, CHEF

gel analysis, column or gradient separation, or other method use in the field to determine and/or analyze DNA molecular weight or sequence content.

II. Methods of detecting and characterizing mini-chromosomes in plant cells or of scoring mini-chromosome performance in plant cells:

Identification of candidate centromere fragments by probing BAC libraries

Brassica oleracea centromere clones are identified from a Bacterial Artificial Chromosome library. Probes are labeled using nick-translation in the presence of radioactively labeled dCTP, dATP, dGTP or dTTP as in, for example, the commercially available Rediprime kit (Amersham) as per the manufacturer's instructions. Other labeling methods familiar to those skilled in the art could be substituted. The libraries are screened and deconvoluted using the protocols recommended by Clemson University (http://www.genome.clemson.edu/protocols/). Genomic clones are screened by probing with small centromere-specific clones (for example 14F1, Figure 1a, was used) which shows high homology to the satellite sequence (14F1 showed homology to "BJCANRD", Genbank ID X68786.1). Other embodiments of this procedure would involve hybridizing a library with other centromere sequences. Of the BAC clones identified using this procedure, a representative set are identified as having high hybridization signals. These are selected, the bacterial clones grown up in cultures and DNA prepared by methods familiar to those skilled in the art such as alkaline lysis. The purified DNAs are fingerprinted by digesting with restriction enzymes such as, but not limited to, HindIII. In a preferred embodiment the restriction enzyme cuts within the tandem centromere satellite repeat (see below). Clones showing a variety of fingerprints are selected for conversion into mini-chromosomes and inheritance testing. It can also be informative to use multiple restriction enzymes for fingerprinting or other enzymes which can cleave DNA.

Finger-printing analysis of BACs and mini-chromosomes

Centromere function may be associated with large tandem arrays of satellite repeats. To assess the composition and architecture of the centromere BACs, the candidate BACs are digested with a restriction enzyme, such as HindIII, which cuts once within the unit repeat of the tandemly repeated centromere satellite. Digestion products

are then separated by agarose gel electrophoresis. BACs containing a large array of tandem repeats will produce a strong band of the unit repeat size, as well as less intense bands at 2x and 3x the unit repeat size, and further multiples of the repeat size. These is prior art on fingerprinting. These methods are well-known and there are many possible variations known to those skilled in the art.

Determining sequence composition of mini-chromosomes by shotgun cloning/sequencing, sequence analysis

To determine the sequence composition of the mini-chromosome, the insert is sequenced. To generate DNA suitable for sequencing mini-chromosomes are fragmented, for example by using a random shearing method (such as sonication, nebulization, etc). Other fragmentation techniques may also be used such as enzymatic digestion. These fragments are then cloned into a plasmid vector and sequenced. The resulting DNA sequence is trimmed of poor-quality sequence and of sequence corresponding to the plasmid vector. The sequence is then compared to the known DNA sequences using an algorithm such as BLAST to search a sequence database such as GenBank.

To determine the consensus of the satellite repeat in the mini-chromosome, the sequences containing satellite repeat are aligned using a DNA sequence alignment program such as ContigExpress from Vector NTI. The sequences are also aligned to Brassica satellite consensus repeats isolated for a previous patent (CrBo1 and 2). The sequences are trimmed to unit repeat length using the consensus as a template. Sequences trimmed from the ends of the alignment are realigned with the consensus and further trimmed until all sequences are at or below the consensus length. The sequences are then aligned with each other. The consensus is determined by the frequency of a specific nucleotide at each position; if the most frequent base is three times more frequent than the next most frequent base, it was considered the consensus.

The previous Chromatin patent (2002) described the determination of consensus. This was taken directly from the literature in human satellite comparisons, and from Hall & Preuss (2002). These methods, including DNA sequencing, assembly, and analysis, are well-known and there are many possible variations known to those skilled in the art. Other

aligment parameters may also be useful such as using more or less stringent definitions of consensus.

Non-selective mini-chromosome mitotic inheritance assays

Assay #1: transient assay

Mini-chromosomes are tested for their ability to become established as chromosomes and their ability to be inherited in mitotic cell divisions. In this assay, mini-chromosomes are delivered to plant cells, in this example *Brassica* suspension cells in liquid culture. The cells used can be at various stages of growth. In this example, a population in which some cells were undergoing division was used. The mini-chromosome is then assessed over the course of several cell divisions, by tracking the presence of a visible marker gene, such as a fluorescent protein. Mini-chromosomes that are inherited well show an initial delivery into many single cells; after several cell divisions, these single cells divide to form clusters of mini-chromosome-containing cells. Other embodiments of this method include delivering mini-chromosomes to other mitotic cell types, including roots and shoot meristems.

Assay #2: Non-lineage based inheritance assays on transformed cells and plants

Mini-chromosome inheritance is assessed on transformed cell lines and plants by following the presence of the mini-chromosome over the course of multiple cell divisions. An initial population of mini-chromosome containing cells is assayed for the presence of the mini-chromosome, by the presence of a marker gene, including but not limited to a fluorescent protein, a colored protein, a protein assayable by histochemical assay, and a gene affecting cell morphology. All nuclei are stained with a DNA-specific dye including but not limited to DAPI, Hoechst 33258, OliGreen, Giemsa YOYO, and TOTO, allowing a determination of the number of cells that do not contain the mini-chromosome. After the initial determination of the percent of cells carrying the mini-chromosome, the cells are allowed to divide over the course of several cell divisions. The number of cell divisions, n, is determined by a method including but not limited to monitoring the change in total weight of cells, and monitoring the change in volume of the cells or by directly counting cells in an aliquot of the culture. After a number of cell divisions, the population

of cells is again assayed for the presence of the mini-chromosome. The loss rate per generation is calculated by the equation:

Loss rate per generation= $1-(F/I)^{1/n}$

The population of mini-chromosome-containing cells may include suspension cells, roots, leaves, meristems, flowers, or any other tissue of transformed plants, or any other cell type containing a mini-chromosome.

These methods are well-known and there are many possible variations known to those skilled in the art; they have been used before with human cells and yeast cells.

Assay #3: Lineage based inheritance assays on transformed cells and plants

Mini-chromosome inheritance is assessed on transformed cell lines and plants by following the presence of the mini-chromosome over the course of multiple cell divisions. In cell types that allow for tracking of cell lineage, including but not limited to root cell files, trichomes, and leaf stomata guard cells, mini-chromosome loss can be per generation does not need to be determined statistically over a population, it can be discerned directly over cell divisions. In other manifestations of this method, cell lineage can be discerned from cell position, or methods including but not limited to the use of histological lineage tracing dyes, and the induction of genetic mosaics in dividing cells.

In one simple example, the two guard cells of the stomata are daughters of a single precursor cell. To assay mini-chromosome inheritance in this cell type, the epidermis of the leaf of a plant containing a mini-chromosome is examined for the presence of the mini-chromosome by the presence of a marker gene, including but not limited to a fluorescent protein, a colored protein, a protein assayable by histochemical assay, and a gene affecting cell morphology. The number of loss events in which one guard cell contains the mini-chromosome (L) and the number of cell divisions in which both guard cells contain the mini-chromosome (B) are counted. The loss rate per cell division is determined as L/(L+B). Other lineage-based cell types are assayed in similar fashion. These methods are well-known and there are many possible variations known to those skilled in the art; they have been used before with yeast cells.

Assay #4: Inheritance assays on transformed cells and plants in the presence of chromosome loss agents

Any of the above three assays can be done in the presence of chromosome loss agents (including but not limited to colchicine, colcemid, caffeine, etopocide, nocodazole, oryzalin, trifluran). It is likely that an autonomous mini-chromosome will prove more susceptible to loss induced by chromosome loss agents; therefore, autonomous mini-chromosomes should show a lower rate of inheritance in the presence of chromosome loss agents. These methods have been used to study chromosome loss in fruit flies and yeast; there are many possible variations known to those skilled in the art..

III. Transformation of plant cells and plant regeneration

Biolistic particle bombardment for delivery of mini-chromosomes

Biolistics can be used for delivery of mini-chromosome into a plant cell, using the standard "dry" or a new modified "wet" delivery method. With the "dry" method, the mini-chromosome DNA-coated microcarriers such as gold are applied onto a macrocarrier and dried. For the "wet" method, the droplet containing the mini-chromosome DNA-coated microcarriers is applied to the bottom part of a filter holder, which is attached to a base which is itself attached to a rupture disk holder used to hold the rupture disk to the helium egress tube for bombardment. The wet biolistics method has been described in detail elsewhere (Mialhe et al., 1995). The concentrations of the various components for coating microcarriers and the physical parameters for delivery can be optimized using procedures known in the art.

Mini-chromosome transformation without selection

Mini-chromosome is delivered to non-regenerable tissues such as *Brassica* suspension cells to obtain stably transformed callus clones for inheritance assay. Suspension cells are maintained in a growth media, for example Murashige and Skoog (MS) liquid medium containing an auxin such as 2,4-dichlorophenoxyacetic acid (2,4-D). Cells are bombarded using a particle bombardment process, such as the helium-driven PDS-1000/He system, and propagated in the same liquid medium to permit the growth of transformed and non-transformed cells. Portions of each bombardment are monitored for

formation of fluorescent clusters, which are isolated by micromanipulation and cultured on solid medium. Clones transformed with mini-chromosome are expanded and manipulated to homogeneity for inheritance assay.

Mini-chromosome transformation with selectable marker gene

Identification of mini-chromosome-transformed cells in bombarded calluses or explants can be facilitated by the use of a selection marker gene. The bombarded tissues are transferred to a medium containing an appropriate selective agent for a particular selectable marker gene. Such a transfer usually occurs between 0 and about 7 days after bombardment. The transfer could also take place any number of days after bombardment.. The amount of selective agent and timing of incorporation of such an agent in selection medium can be optimized by using procedures known in the art. Selection inhibits the growth of non-transformed cells, thus providing an advantage to the growth of transformed cells, which can be further monitered by tracking the presence of a fluorescent marker gene or by the appearance of transformed explants (transformed cells on explants turned green under light in selection medium). These transformed cells can form shoots directly, or alternatively, can be isolated and expanded for regeneration of multiple shoots transgenic for mini-chromosome.

Useful selectable marker genes are well known in the art and include, for example, herbicide and antibiotic resistance genes including but not limited to neomycin phosphotransferase II (conferring resistance to kanamycin, paramomycin and G418), hygromycin phosphotransferase (conferring resistance to hygromycin), 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS, conferring resistance to glyphosate), phosphinothricin acetyltransferase (conferring resistance to phosphinothricin/bialophos), MerA (conferring resistance to mercuric ions). This is similar to ordinary transformation; there are many variations possible

Regeneration of *Brassica* plants from explants to mature, rooted plants

Regeneration of a whole *Brassica* plant involves culturing of regenerable explant tissues taken from sterile seedlings or mature plants on a shoot regeneration medium for shoot organogenesis, and rooting of the regenerated shoots in a rooting medium to obtain

intact whole plants with a fully developed root system. These plants are potted in soil and grown to maturity in a greenhouse.

Explants are obtained from any of the tissues of a *Brassica* plant. These tissues include hypocotyls, internodes, roots, cotyledons, petioles, cotyledonary petioles, leaves and peduncles, prepared from sterile seedlings or mature plants. Further, the *Brassica* tissue can be from any *Brassica* species such as *Brassica* napus, *Brassica* oleracea, *Brassica* nigra, *Brassica* carinata, *Brassica* juncea, and *Brassica* campestris.

Explants are wounded (for example with a scalpel or razor blade) and cultured on a shoot regeneration medium (SRM) containing Murashige and Skoog (MS) medium as well as a cytokinin, e.g., 6-benzylaminopurine (BA), and an auxin, e.g., α-naphthaleneacetic acid (NAA), and an anti-ethylene agent, e.g., silver nitrate ($AgNO_3$). For example, 2 mg/L of BA, 0.05 mg/L of NAA, and 2 mg/L of $AgNO_3$ can be added to MS medium for shoot organogenesis. The most efficient shoot regeneration is obtained from longitudinal sections of internode explants.

Shoots regenerated via organogensis are rooted in a MS medium containing low concentration of an auxin such as NAA. Plants are potted and grown in a greenhouse to sexual maturity for seed harvest.

To regenerate a whole plant transformed with a mini-chromosome, explants are pre-incubated for 1 to 7 days (or longer) on the shoot regeneration medium prior to bombardment with mini-chromosome (see below). Following bombardment, explants are incubated on the same shoot regeneration medium for a recovery period up to 7 days (or longer), followed by selection for transformed shoots or clusters on the same medium but with a selective agent appropriate for a particular selectable marker gene (see below).

Figure 7A is a picture of a regenerated brassica plant transformed with a mini-chromosome. Figure 7B is a micrograph of a portion of a leaf of the same plant, showing the edge of the leaf. Figure 7C is a micrograph of a portion of a root of the same plant, showing the tip of the root. Figure 7D is a fluorescence image of the same leaf as in B, showing expression of the fluorescent protein gene encoded by the mini-chromosome. Figure 7E is a fluorescence image of the same root as in B, showing expression of the fluorescent protein gene encoded by the mini-chromosome.

Assays of mini-chromosome presence in transformed plant tissues

Mini-chromosome presence can be assayed for in transformed cells and tissues by methods including but not limited to examining the tissue for the presence of a fluorescent gene product from a marker gene present on the mini-chromosome, examining the cells for the presence of a pigmented gene product from a marker gene present on the mini-chromosome, examining the cells for a visible or fluorescent enzymatic product of one or more genes present on the mini-chromosome, examining the cells for binding of a fluorescent or otherwise visible protein to a specific DNA sequence on the mini-chromosome, Fluorescence In Situ Hybridization to specific DNA sequences on the mini-chromosome, or In Situ Hybridization to RNAs produced specifically from the mini-chromosome. These are all similar to other published methods, there are many possible variations.

Determination of mini-chromosome structure in transformed plants and tissues

The structure of the mini-chromosome in transformed plants and tissues can be determined by methods including but not limited to: Southern blot hybridization to genomic DNA from transformed tissue subjected or not subjected to restriction endonuclease digestion, dot blot hybridization of genomic DNA from transformed tissue hybridized with different mini-chromosome specific sequences, PCR on DNA from transformed tissues with probes specific to the mini-chromosome, or Fluorescence In Situ Hybridization to nuclei of transformed cells.

Furthermore, mini-chromosome structure can be determined by DNA sequence analysis, or other methods of examination of mini-chromosomes 'rescued' from transformed cells. In the mini-chromosome rescue assays, genomic DNA from transformed cells is transformed into laboratory E. coli strains. The E. coli are then selected for resistance to a particular antibiotic whose resistance gene is present on the mini-chromosome. Circular mini-chromosomes are able to replicate in E.coli and contain specific bacterial antibiotic resistance genes. Once resistant bacterial strains have been identified, the mini-chromosome DNA can be purified from the bacteria and subjected to various assays, including but not limited to DNA sequence analysis, restriction

endonuclease digestion, agarose gel electrophoresis, CHEF gel electrophoresis, and Polymerase chain reaction.

Mini-chromosome autonomy demonstration by In Situ Hybridization (ISH)

We can assess whether the mini-chromosome is autonomous from the native plant chromosomes, or has integrated into the plant genome by In Situ Hybridization (Fluorescent In Situ Hybridization or FISH is particularly well suited to this purpose). In this assay, we take mitotic or meiotic tissue, such as root tips or meiocytes from the anther, possibly treated with metaphase arrest agents such as colchicine, and use standard FISH methods to label both the Brassica centromere (using probes from sequence 14F1, Figure 1a), which labels all Brassica chromosomes with one fluorescent tag (Molecular Probes Alexafluor 568, for example), and label sequences specific to the mini-chromosome with another fluorescent tag (Alexafluor 488, for example). All centromere sequences are detected with the first tag; only mini-chromosomes are detected with both the first and second tag. Chromosomes are stained with a DNA-specific dye including but not limited to DAPI, Hoechst 33258, OliGreen, Giemsa YOYO, and TOTO. An autonomous mini-chromosome is visualized as a staining body separate from the native chromosomes.

Determination of gene expression levels

The expression level of any gene present on the mini-chromosome can be determined by methods including but not limited to one of the following. The mRNA level of the gene can be determined by Northern Blot hybridization, Reverse Transcriptase- Polymerase Chain Reaction, binding levels of a specific RNA-binding protein, in situ hybridization, or dot blot hybridization.

The protein level of the gene product can be determined by Western blot hybridization, Enzyme-Linked Immunosorbant Assay (ELISA), fluorescent quantitation of a fluorescent gene product, enzymatic quantitation of an enzymatic gene product, immunohistochemical quantitation, or spectroscopic quantitation of a gene product that absorbs a specific wavelength of light.

Mini-chromosome rescue from transformed plant cells:

Circular mini-chromosomes that contain bacterial sequences for their selection and propagation in bacteria can be rescued from a transformed plant or plant cell and re-introduced into bacteria. If no loss of sequences has occurred during replication of the mini-chromosome in plant cells, the mini-chromosome is able to replicate in bacteria and confer antibiotic resistance. Total genomic DNA is isolated from the transformed plant cells by any method for DNA isolation known to those skilled in the art, including but not limited to: rupturing of plant cells by freezing, grinding the frozen tissue to a fine powder, extracting DNA with cetyl-triethylammonium bromide (CTAB), removing debris by centrifugation and extracting several times with organic solvents, (i.e. phenol, chrloroform, isoamyl alcohol or mixtures thereof) to purify the DNA, followed by ethanol precipitation of the extracted DNA. The purified genomic DNA is introduced into bacteria (i.e. E. coli) using methods familiar to one skilled in the art (for example heat shock or electroporation). The trasnsformed bacteria are plated on solid medium containing antibiotics to select bacterial clones transformed with mini-chromosome DNA. Transformed bacterial clones are grown up, the plasmid DNA purified (by alkaline lysis for example), and DNA analyzed by restriction enzyme digestion and gel electrophoresis or by sequencing. Because plant-methylated DNA containing methylcytosine residues will be degraded by wild-type strains of E. coli, bacterial strains (i.e. DH10B) deficient in the genes encoding methylation restriction nucleases (i.e. the mcr and mrr gene loci in E. coli) are best suited for this type of analysis. Mini-chromosome rescue can be performed on any plant tissue or clone of plant cells transformed with a mini-chromosome.

Use of exonuclease to isolate circular mini-chromosome DNA from genomic DNA:

BAC-end sequencing procedures, known to those skilled in the art, can be applied to characterize mini-chromosome clones for a variety of purposes, such as structural characterization, determination of sequence content, and determination of the precise sequence at a unique site on the chromosome (for example the specific sequence signature found at the junction between a centromere fragment and the vector sequences). In particular, this method is useful to prove the relationship between a parental mini-

chromosome and the mini-chromosomes descended from it and isolated from plant cells by mini-chromosome rescue, described above.

Structural analysis of mini-chromosomes by BAC-end sequencing:

BAC-end sequencing procedures, known to those skilled in the art, can be applied to characterize mini-chromosome clones for a variety of purposes, such as structural characterization, determination of sequence content, and determination of the precise sequence at a unique site on the chromosome (for example the specific sequence signature found at the junction between a centromere fragment and the vector sequences). In particular, this method is useful to prove the relationship between a parental mini-chromosome and the mini-chromosomes descended from it and isolated from plant cells by mini-chromosome rescue, described above. This is a routine procedure.

Methods for scoring meoitic mini-chromosome inheritance

A variety of methods can be used to assess the efficiency of meiotic mini-chromosome transmission. In one embodiment of the method, gene expression of genes encoded by the mini-chromosome (marker genes or non-marker genes) can be scored by visible methods (i.e. fluorescence of fluorescent protein markers, scoring of visible phenotypes of the plant), by scoring resistance of the plant or plant tissues to antibiotics, herbicides or other selective agents, by measuring enzyme activity of proteins encoded by the mini-chromosome, or by measuring non-visible plant phenotypes, or by directly measuring the RNA and protein products of gene expression using microarray, northern blots, in situ hybridization, dot blot hybridization, RT-PCR, western blots, immunoprecipitation, Enzyme-Linked Immunosorbant Assay (ELISA), immunofluorescence and radio-immunoassays (RIA). Gene expression can be scored in the post-meiotic stages of microspore, pollen, pollen tube or female gametophyte, or the post-zygotic stages such as embryo, seed, or progeny seedlings and plants. In another embodiment of the method, the mini-chromosome can de directly detected or visualized in post-meiotic, zygotic, embryonal or other cells in by fluorescence in situ hybridization, in situ PCR, PCR, southern blot, or by mini-chromosome rescue described above. This is published before.

ISH analysis of mini-chromosome copy number in meiocytes, roots or other tissues of transformed plants

We can assess the copy number of the mini-chromosome in any cell or plant tissue by In Situ Hybridization (Fluorescent In Situ Hybridization or FISH is particularly well suited to this purpose FISH). In this assay, we use standard FISH methods to label both the Brassica centromere (using probes from sequence 14F1, Figure 1a), which labels all Brassica chromosomes with one fluorescent tag (Molecular Probes Alexafluor 568, for example), and label sequences specific to the mini-chromosome with another fluorescent tag (Alexafluor 488, for example). All centromere sequences are detected with the first tag; only mini-chromosomes are detected with both the first and second tag. Nuclei are stained with a DNA-specific dye including but not limited to DAPI, Hoechst 33258, OliGreen, Giemsa YOYO, and TOTO. Mini-chromosome copy number is determined by counting the number of fluorescent foci that label with both tags.

Induction of callus and roots from transformed plants tissues for inheritance assays

Mini-chromosome inheritance is assessed using callus and roots induced from transformed plants. To induce roots and callus, tissues such as leaf pieces are prepared from transformed plants and cultured on a Murashige and Skoog (MS) medium containing a cytokinin, e.g., 6-benzylaminopurine (BA), and an auxin, e.g., α-naphthaleneacetic acid (NAA). Any tissue of a transformed plant can be used for callus and root induction, and the medium recipe for tissue culture can be optimized using procedures known in the art. This is published.

Clonal propagation of transformed plants

To produce multiple clones of plants from a mini-chromosome-transformed plant, any tissue of the plant can be tissue-cultured for shoot organogenesis using regeneration procedures described under the section regeneration of *Brassica* plants from explants to mature, rooted plants (see above). Alternatively, multiple auxiliary buds can induced from a mini-chromosome-transformed plant by excising the shoot tip, which can be rooted for a whole plant; each auxiliary bud can be rooted for a whole plant.

Scoring of antibiotic- or herbicide resistance in seedlings and plants (progeny of self- and out-crossed transformants

Progeny seeds harvested from mini-chromosome-transformed plants can be scored for antibiotic- or herbicide resistance by seed germination under sterile conditions on a growth media (for example Murashige and Skoog (MS) medium) containing an appropriate selective agent for a particular selectable marker gene. Only seeds containing the mini-chromosome can germinate on the medium and further grow and develop into whole plants. Alternatively, seeds can be germinated in soil, and the germinating seedlings can then be sprayed with a selective agent appropriate for a selectable marker gene. Seedlings that do not contain mini-chromosome can't grow and die; only seedlings containing mini-chromosome can survive and develop into mature plants.

Genetic methods for analyzing mini-chromosome performance:

Fertile plants transformed with mini-chromosomes can be crossed to other plant lines or plant varieties to study mini-chromosome performance and inheritance. In the first embodiment of this method, pollen from a transformed plant can be used to fertilize the stigma of a non-transformed plant. Mini-chromosome presence is scored in the progeny of this cross using the methods outlines in the preceding section. In the second embodiment, the reciprocal cross is performed by using pollen from a non-transformed plant to fertilize the flowers of a transformed plant. The rate of mini-chromosome inheritance in both crosses can be used to establish the frequencies of meiotic interitance in male and female meiosis. In the third embodiment of this method, the progeny of one of the crosses just described are back-crossed to the non-transformed parental line, and the progeny of this second cross are scored for the presence of genetic markers in the plant's natural chromosomes as well as the mini-chromosome. Scoring of a sufficient marker set against a sufficiently large set of progeny allows the determination of linkage or co-segregation of the mini-chromosome to specific chromosomes or chromosomal loci in the plant's genome. Genetic crosses performed for testing genetic linkage can be done an a variety of combinations of parental lines; such variations of the methods described are known to those skilled in the art.

It should be understood that various changes and modifications to the presently preferred embodiments described herein will be apparent to those skilled in the art. Such changes and modifications can be made without departing from the spirit and scope of the present invention and without diminishing its intended advantages. It is therefore intended that such changes and modifications be covered by the appended claims.

CLAIMS

The invention is claimed as follows:

1. A *Brassica* centromere composition comprising an isolated sequence selected from the group consisting of SEQ ID NO:1, NO:2 and SEQ ID NO:3 or a fragment thereof.

2. An artificial plant mini-chromosome comprising one or more plant-expressed genes, the gene comprising a promoter, a coding region, and a terminator sequence separated by restriction endonuclease sites or recombinant sites, where in the promoter is a non-plant promoter derived from a non-plant species.

3. The artificial plant mini-chromosome of claim 2, wherein the gene comprises an intron.

4. The artificial plant mini-chromosome of claim 2 further comprising a bacterial plasmid backbone.

5. The artificial plant mini-chromosome of claim 4, wherein the bacterial plasmid backbone is selected from the group consisting of a replicon of F' plasmidof E.coli, a bacteriophage P1 replicon, and an antibiotic-resistance gene.

6. The artificial plant mini-chromosome of claim 2 further comprising a telomere.

7. The artificial plant mini-chromosome of claim 2 further comprising a centromere.

8. The artificial plant mini-chromosome of claim 7, wherein the centromere is the centromere of Claim 1.

9. The artificial plant mini-chromosome of claim 2, wherein the mini-chromosome is circular and the chromosome is without a telomere.

10. The artificial plant mini-chromosome of claim 2, wherein the mini-chromosome is linear and the chromosome has a telomere.

11. The artificial plant mini-chromosome of claim 2, wherein the mini-chromosome does not contain any bacterial sequences required to propagate the mini-chromosome in bacteria.

12. The artificial plant mini-chromosome of claim 2 wherein the promoter is derived from a non-plant species.

13. The artificial plant mini-chromosome of claim 12, wherein the non-plant promoter is selected from the group consisting of SEQ ID NOS:4 to 23.

14. The artificial plant mini-chromosome of claim 2, wherein the mini-chromosome is derived from a donor clone or a centromere clone having substitutions, deletions, insertions, duplications or arrangements of one or more nucleotides in the mini-chromosome.

15. The artificial plant mini-chromosome of Claim 14, wherein the mini-chromosome is obtained by passage of the mini-chromosome through various hosts.

16. The artificial plant mini-chromosome of Claim 15, wherein the host is selected from the group consisting of viruses, bacteria, yeasts, plants, prokaryotic organisms, and eukaryotic organisms.

17. A promoter for expressing genes in plants wherein the promoter is derived from a non-plant species.

18. The promoter of Claim 14, wherein the promoter is selected from a group consisting of SEQ ID NOS:4 to 23.

19. A transgenic plant comprising the mini-chromosome of Claim 2.

20. The transgenic plant of Claim 19, wherein the plant is a *Brassica* plant.

ABSTRACT OF THE DISCLOSURE

The invention is generally related to methods of generating plants transformed with novel autonomous mini-chromosomes. Mini-chromosomes with novel compositions and structures are used to transform plants cells which are in turn used to generate the plant. Methods for generating the plant include methods for delivering the mini-chromosome into plant cell to transform the cell, methods for selecting the transformed cell, and methods for isolating plants transformed with the mini-chromosome. Plants generated in the present invention contain novel genes introduced into their genome by integration into existing chromosomes.

Figure 1, sequences relevant to Brassica mini-chromosome BB5R4-1

1a. Sequence of clones 14F1, used to screen Brassica BAC library for candidate centromere clones (SEQ ID NO:1)

ACCCTNCCCAAACTGGGAAACTGGAATCACCTGATTTGAAAGTGGGATAACT
TCTTCATGCCAACTCCTATGAGTTTTATTCAACTTCCTGGTGATTCTCCACCAC
TTTATGTATCCAAATCAAGCTTCTTACAAAGTGATTCATCCTGGTTTGATTGG
AACGACGAACAAGTTGTGCTATTCCCAAACTTGGAAACTGGAATCACCTGAC
TTGAAAGTGGGATAACTTCTTCATCCCAACTCCTATGAGATTTATTCAACTTC
CTGGTGATTCTCCACCACTTTATGTATCCAAATCAAGCTTCTTACAAAGTGAT
TCATTCTGGTTTGTTTGGAACGACGAAGAAGCGGGGATCCTCTAGAGTCGAC
CTGCAGGCATGCAAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTTGGCGT
AATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCA
CACAACATA

1b. Sequence consensus for all centromere satellite repeats from BB5R4-1 (SEQ ID NO:2)

ASCTTSATTTGGATACATAAAGTAGTGGAGAATCACCAGGAAGTTGAATAAA
TCTCATAGGAGTTAGGATGAAGAAGTTATCCCACTTTCAAATAAGGTGATCC
CAGTTTYCCTGTTTGGGAATATKANAACTTHTTCGHCATTCTADTCAAACCAG
GATGAATCGCGATGTAARARVCY

1c. Unique sequence present in BB5R4-1 (SEQ ID NO:3)

TTGATCTCTTCAACTCAAACACACGGCTCAGATTAGAGATGTTACCGTAGACC
TTTTGGAGCGTATCCCACAGGTCCTTTGGCGTCTCACAGTAGCTGTAGGCTTC
CAGGATTGAAGCTTCAAGTGACCCATGTAGTACAGTAAGCACCTTCAAGTCA
TCTTGATCCCACTTCTCTTGATCTACCACCATCAGCTCTTGACCGCCTTCTCCT
TCTTTAGCAACTGGTTTCGGAGCTCCATCTGAGATATGGCTCCATAAACCCTT
GCTCCCAACTGCCGCTTTCACCAAGCGAGACCACAACAGGTAGTTAGTACCA
CCCTTCAACTTAACCGCAACCGAGAGAGCCTTGCTCTTGCTTTCTTCCATTCC
AAATTCTTAAGAACAACCGTAAATGAAATGAAGCTGAAGATTACTGAGATTT
AGAAACCAGGAAATGGGTAACCTGGCTCTGATACCATGAAAGATTAAGAGA
ATGAAATGATTTAGAGAAGATTAAGGAACAGAATGAGAGATTATGAAGAGA
TTAGACTAGAATCATGATTAAGAGAGTATGAACTTAGAGAGATAAACTCAGA
ACTGTATGGTTTATTATTAATGACAAGAGTTACAATATATAGTGTGAGATCAT
AAGGGTTTTCGAAATCAAAGAGCTAAGCTAAG

Figure 2: Non-plant derivative promoter sequences

SEQ ID NO:4 (5' to 3')

gttgtccgcagcggagatgcaactgatgcaacccacatttcagatcaccgacaacgtgcagcgcggcaactacgccactctgaccgacaaggatgtggcgcatttcgagcagctcctgggcaaga
acttcgtgctcactgaggacctggagggatacaacatctgcttccttaagaggattcgaggtaggttgtgtaaccaaattcattcacattcgtgtgcccttaatgaatttctccgatgaattgcttcaacca
ggcaacagcaagttggtgcttaagcccggaagcacggcggaggtggccgccatcctgaagtactgcaacgagcgtcgtttggcggtggtgccgcagggcgggaacacaggtctagtgggcgg
atccgtgccgatctgcgacgagattgtccttctctagcgcgcctgaacaaggtgttatccgtggacgaggtcaccggcattgctgtcgtggaggcgggctgcatcctggagaacttcgatcagagg
gccagagaggtgggcttgacggtgccactggacctgggcgccaaggccagttgccacatcgggggcaatgtgtccacaaacgcgggcggagtgcgggtggtgcgttacggcaatctgcacgg
ctctgttttgggcgtggaggcggtgctggccaccggtcaggtgctggaccttatgtccaacttcaagaaggacaacaccggctaccacatgaagcacttgttcataggatccgagggcactctgggc
gtggtcacgaagctttcgatgctctgcccccattcctcgcgagcggtgaacgtggccttcatcggcctgaactccttcgacgatgtgctgaagacttttgtcagtgccaagcgtaatctgggcgagatt
ctaagctcctgcgagctgattgacgagcgggccttgaacaccgccctcgagcagttcaagttcctgaagtgagttgcgccacctttgtcttctctgagcgttaccaatcctgttcacaaacttatttccca
tagctcccccatttcgggatttcccttctacatgctcatcgagacctcgggcagcaacggtgaccacgacgaggagaagatcaaccagttcattggggacggtatggagcgtggcgagatccagga
tggcaccgtaaccggtgatcccggcaaggtgcaggagatctggaagatccgcgaaatggtgccgctgggtctgatcgagaagagcttctgcttcaagtacgacatctcgctgcctctgcgggactt
ctacaacattgtggacgtgatgcgagagaggtgcggtcccctggccacagttgtctgcggatacggccatctgggggactctaatctgcacctgaacgtctcctgcgaggagtttaacggcgagatc
tacaagcgggtcgaaccttcgtctacgagtacacctccaagctgaagggcagcattagtgcggagcacggcattggcttcctgaagaaggactacctgcactactccaaggacccggtggccatt
ggctacatgcgcgagatgaagaagctgctggaccccaacagcatcctcaatccctataaggtgcttaactgaaggcttctacctaatagattctatttttttttgtttgtgtgtaattttcataaccttataatac
agaaatggcattagaagtgaatttttgttaacttgtgaagttaaaaaggaccatcatatttggcacgaaaccaatgggcaaaacttacttataaaaatagtccgaaaaaatagtatatatccagttttacagta
ccacattataggtactcggaggtaataatagaaaaaacactatctttgcatttactgttacactacgaagcactatatttagtagcagtactcattagagtccactcacaaaattagcaccaaccggcagta
attggtcaaggatcggcgatagcttcaaactccgaagttcaaagtcaaactgccgccctgcgaaagcttcgcgagtggagcttttctgcacttatcgatagctaacattgtggcgcgactatcgatcga
cgagctgccgcttaacagtgccatatatagattgtaacattagaagctcaaatcattgttggagcacaaaccacaaagaacacacgaaac

SEQ ID NO:5 (5' to 3')

aaaatatttcacctcattttccgcacaccatttataagcaaagttacccccaacccataacttttatggtaagtaatacagaccctccaagttcggcaaatcgatacccagcgaccttgagcttgacatttat
atatatgccagaatataacgaccacgtgctgtcaactgtgtcaggaaaagctcacccacacttttctttggaggagctgtgctccctaaacgaatttcattgtcaaggtcgcacgcacaaaaatgaagag
gaaaagctgaatgtgggtggaaatgccggccggcacgaccttgaagccagttgggtgagaaataaaaagcttttgccggtaggagacttgtggaacatcacccacaagtggcggacttggccttg
gcgatggccttgttggagctccctcagcaaaaatgttacatagggggaggaaataagctcaattggctttatgctttccgctccctggaagtccttttctggaatgttaaagtgttaaatgacatttattgaa
catttgggacagaggaggagataatacaatatacttgtctaatttaaaaaaaatcgttattatgatttattccatatgtaagattttaattcatcatgattgtaaataaattatatataaacaaattcaataaatttac
attattgataaaatttatttttttcatgaaattatacccaaaaattattctcaatttttcttataatcagttttgcataagtatactttcttcataccccctctaccacagccactgctttcttgactttgcaactatccggg
aacagcttatcataatggatgagctgcagctaacggaaaatgggggagctgggatcaaacatttttccaaggttgaaattgtcgtcagcataatgtttgagggagctggattcgcgttagcttgaaggtc
aatccatttgggtgcccttgttatggtcaagtttaaggctgcaatagggggaatcttcaaggaccattacgcaaggttttcgcatcaaagatttgccgtgcaagctttttgagttgaaggatgcttaacttg
aaagcgggttagtggttccaagagattttaggtgaaggagactccgctgttttgaaatatattaagtatgtaaagaagtatactataaataacccaaagtgatacaatgtaagaaaagatctcgttggtcc
ctggtataaatttgtttgccattaatgaatattgaaaataataattatactaataataggtacaataagcaagattaaattgcatttaatcaccaaaaatcagtttctatgcgaaccaaaatgtcataacaaaca
attgttgattcatccgtagtgaaatccaagttcgaaattcgaaatgagcatacgacgaccaaacttcccctcaaaattgctagactcagctagagcaagtacgcccaagttaaccccctgaaattcgaaat
gaattcgatgccgcgcttcgaacaacgaaatcccaaagagcttacgttttatttgacgtagcactcttacgtgaaatgattttccccaattccgctctcatttcccgagtctctcaccgcttctcagccacttt
cccaccccctttctagttccgaagtaaaggtaacaaaggcagccgtgtctttggggtggtaaactggcggtggtggtggcacattgtcagtggtgtgggttcctgtggttggtggttcaattggttggtt
gttggcataaacaaagcacacacacaatacacacaaactcccgggggggtggtggaaattgggagggtgacattcactgcgagagaggaactcgcttcctataggaaagtacaaagagagctatttt
ataaatgtgactgcagcaaggatatttacagtcagtccactctgaaacctcgacgagagaacattgaataacaagcggaagcgaaaagcgcagttgaaagttcgtcaaaaagcgacaagtttcctcg
ttcgtttcccgccaaatgagtcagaaaaattttccaagtgctcgatacgaaacataaagacttacaagacttaaagtgcaagcagtgaatggaatatattattcctcagcgatattgaaatcaaacattaa
aaatatatgctacactaaagttatatatttttttaaagattcatacgttttgtaaaatcacattttgtattaaattaaataccgcc

SEQ ID NO:6 (5' to 3')

tgggtgcgtcgcaggtttcactggaaaacaatttgcacttttgtttgtggagtcgacaacaaaagcattcacttgtctaagactctctcattcataactcgcactttagttcactgaaccgcacgcaaaactt
tggggcggacaacatgttttcgaggtgccaaaagcttcataaaactaccaatccattagattaaattccaggcggtacatcttttggggatgattcatgtggcaggggttctctactcgtttacaatcatat
catcatcttcaagatcatatagtttatcatatcagtagagtactacaatataatgcataaactaagccaaataactttatgacgcgtgcttatgcgaaagtaaactttattatcaaatttacttaaccgtgaaatc
aaaaccttttatataaacacgaatattattatctttgctaaataaaactctcgcttaacaaacaatgacacttcaattccaacatagagtttatcttaagccaataaccaaaaacggaacttacataacttgcca
acaaacatatgaatatagctatttcggatcgtgggagaccattatgcatacaaggcacgctcctaaaaaccgtgttaaacaaatatatgtcaaatgtatatcttaaaaaagcgcgcacatatcttttgaaat
atcttcacccagagtatgtatgagattaaactggattagcactaagccacagcttctgtagatagaaatttatgcagagagtagattatttggctgctgagcaatttgaccaccacaagatagcagagaa
catctgacattttctatatccatataataaaactgacttaacactaagctgaagtggtatgtttaaatcctccagctaataaatcgagactaaacgccctatcttatagtgatatataatagtatctatatgtgtat
tgtcatttactgtttatgagtatttgaaaaaaccattctatatttttataggttagttaataaatattttgatatacatatgtagattggctcacacgtacttatgacccactacataataaaattgttttgttttttaatag
aataatggtttataaaaagtttagactcacacggaaatgataaactctttgcaaatacagctttcatttttattacaaattgcactctttcagatctgcagttgctatgccaacctttttattccctttactaaaaggg
tatactaggcttactgaacagtatgtaactggtaaagtaaagcgtttccgattctataaattatatatctaaacttttgatcagtcgaatccatctgaacacattctgtcacattagattattccagaaactcaac
ttaaacatgtgtatttttttaagaccattatcaaggatattaaaaatggtctcctaaaatttaataaacaaaagtgtcacatcaaatttaagacgtaaattaatatttttttttctatggtgaaataattgttatttttccaa
tgttgtgaaataataaatgtatcttttcaacgcacacattttcaaggttttaataataatagtgactcgtgcgtgaataagagagaaattaagattttaaaaaagaataaaattcagagatgtgatctgtaaaa
attatttaccaatttttcatttacccccgaaagtgatgctaatggttaaaacggcatttgcgacttatctcctacgtaatattgcaaaaataaggatttggttagatgagtgtgaagtaaacaagatgcaaagtt
ttggagatagaaaacatagccttgagtcttggtcatgtttacttggcaccaggccgcgattatcagcgctactagtcgtaatttgagttagaccttaatactctaagtgagagtgatgatatacgatttccc
agccacttgctttctacgaaatgcgctaaaaaaaatccctaactacacaaagatttgtgttgttatccaggtgttctgatataaaaggcggcaaggaaattgatggcatcatcagtatcaaagtgagagtg
attgcagtcacac

SEQ ID NO:7 (5' to 3')

atgggacggtcctattctcagcaaaaattgacaagaacaacaacaatgtctatggaaaatcgaacttcatcccagcacctgcagaaatcccgagcgagtcggggaaaaagtatttaacccccgaaa
gggttttccccaaaataatgaagtaatgaatgaagcggaaaacactggccgccaatctacctaatactaatgagcgggccaacccgaccaggaatttttgcaagtcaggtacttcaacggatatatgg
gttcgacaagtgcggattttcccgcgacatcaatgaggacttggccgggttatccgcggtgctcatcgggcaattccgcggccgaggacttcatcgtagtgatcattaggtagatatgtgcatggatgt
gacatggcgatcattgcgcggaataacacacgtaataaccgagatatccgggatgacccaccaggtaggatgtgaggacatatagaaaacccccagccagttttccactcgtcgtggcttgttttgc
ttgagtttcgctgactgcgtaattggataagatgggaaattacttttaaatccttcgctgatccacatccggacattcgtcgaaggaaaatccattgcagggaaatacgaaatggaaatgcggctgggtta
ttggctcgacatttcccatcttccctcacgccattggttgcaggatcgcggggaattggaattccgcgctggaatttttgtcacctcttgggtttatcaaaacttttgggtttgctatggatttttccaatttta
ccaccgcgcctggttttttttttttgacgacgcggaaaatcggacttggctatgcgggcttgtctgtttttccgggtacaaagtctgcatgtcagcctccatgcgggagtgggagttgggaaagtttcccat

cgatagttggaggggtggcttgaaagtctggaggtgctagctgggaaagttgtgtgtgcgcgatgaggcaaggagtcaaagatcaggggagttggaaagcgagaattgtgggaatcgtccagga
ctcagctggatgctgaggggcagtatgatttttttacgttatcaatcgaattgattttaagacagcagaacttcacatactaataagatgaccatgggattagttaaaatgtgtaactcgtattcgaatcgtc
attctttcacggaccaatcgtgggaacaggagatctcttcgatccaagctcacaggagacttgacactcttcgtctattccttgtcaagtttttaatgacatctcctatgccctgagctatgttttcctagctct
catcgatcgctgccaatgagccactggagatgatccataagtcagcgtagagtgcaccccagagttgacacttggtgtctcggaattcggctcattatcagtgctatttttggaacacctctctgcgaag
gtgtcatttttgtcagtgcgtatcgctcaggttcaactccccaccaaaaaccgaatttagagcatcggcagatgtacttgaagcactcaatctaagtgaggaaaccaccccatgaacgaagagtactag
gagtcctatttgactcgtgcttaaaaatagaaaattacttagggtgatccataggtagggaggcgatattgtaacttgcatttcggacccggacctgcacgagttattacgggtgggttgtgagcgtatcg
ggaaattggagagccaccagatctgtcataacttatacgggggatccttattcctgggagggtgcgcctgcgtctgctcttccgagagagaggtgggaaatggaggaagagagagagagagaga
gtgagagagcaggtagagggaagtgagggaaatacgcaataagggtatgggaaaagtgctgttgttgttgctaggtagcgacgcacacgtgcgagtgtttttctgttttgaagaagaaccaccacca
aatggcgacagcggcgtcggcagaggcgcagagttccgggtataaaagagcgtgctcgactgttgacctgtcacagccacctcagctctcgttgagaacgcaaccaccgctctatactcgatccc
gaactatataactcgcctctcgatcgccgatctcccgatttacccatctcgatcagtaccggaaacc

SEQ ID NO:8 (5' to 3')

atttggctccccatcgccatcggttgctccaatgacactagggaattgtgggccgccgacagctgtccttaattacatggaaatccacactagattcgtgcccctcgccccgtactcgcagccgaagtc
cccacagagtcattcaccttgccaccaccaaaaaaaaaacgaaagcaactgaaggaaaagttcgattcgaaggctgagggataccccttaaaggcccatttcccggcttcgtaaatcacatttagttag
ccatttagactacagcaagtcttttaagatacactgcaaaataaataccattacattaatagaagtgtcatgtcatcggtctgtatttttgttaccacagaatagacttacatatatgataaaaaaatgttcaac
aataagttacatcggtagccaattctatagatttaattccttacgaatatagtttcgttggaatactcaatttgtaattgtaattaattataattattataattttaagaatttatataagtaactaaaagacacggca
gacacagaatgaaaacactctatgttagggaatgcaaaaaaacgtggcggaagccaaaaggcgcaagcaaaaatcgaaaccaagtgaatatacatattatttcaacaggcaactcattcagcatat
aatattaccacccatggagctttatgtagttgatgtacgtagtctatgatgtggagcccacgttggcggaactgggaatggggattggggtttgagagctgtggtaaattgggggggttgaagtatcaag
ggtttgggttctgtagacctgcggaatcgaggtgaataagcgaagaacacattcacacacactaaaaggcaaacaaagggaaatcaatctttgtacatacttttagcatatgcacacgtatgatctcca
cccacttttccctcccaatgaaacaaacacacacacacatgcaaggccgtacgtttgtatatgtgtgcggttgtcggctttgccgggaattggggaatatttgcatgcctttgtgtacttttccatatgattt
atgacctaaattgttgctgctcgcgcacatataattacacacacatcgctgtggccatgtgtgtgtgtgtcgtcttgggacgcgcgccaaagtatgctacacttttgttttatgagttaataagtaggcgtg
gccccagcccaattgctacactctgattatggcaccggatacccagatagacgcccatccaccccactgtaagatgggggaatttccaaacctatatgtatgtgcagatcagatagatagcacaga
acttttaaagtacacttttggggcacgcaatttagaaaatgtacctcggtgtcggagaaattattttaaaagtcgactgaaccacctcgttccatatggagaagtctacgagttcaagtttaatggagcag
ctgactgcactgaattttgtagtttaatacacaaatccgcaaattgcatctcacttcaaatagcctggtacatagtatctactaacataactcatattaaaataaagcaaccaaccagagggccgaagttct
attaataaaactaatatttaactattatatatacatttttatttacttggtacgcttatgataaccttcgaaagagaaccaacacaatacgctttgtcatttgaaaaataaatatgctgtaactactttacaaggtga
aactcttgtcagaagataagaggctaggtaagttgattattcaatcagtttacttactgcaacccaaaatggtcactgcactaaccttcagatgagctgcactacaccctcaatcgagaatcaatgcaaac
gcagtgccagcgaaaatgtcagcaagggattaggccaatcccaaacgggtaatcccgctgcgacaatgctaatccaattccgatgggccgtataaaagccccaagctgggctggctgtgatttcgt
cttggcccgcagaccggagcatggagtccggtaacgtgtcgtcgagc

SEQ ID NO:9 (5' to 3')

atcgatgacggcatcggcttgacctctcggagtacgtttgattttatagaacaagttttctccttcttatactataaggaaaaattataaaaattgctgaaaatgaaacatggctagaattcgtttttaacattt
tttcaatctgagaaaaaatttccgattagtcttaaaataactaaaccaattcgtataccgttaatcgtagaagaaaaatgaaattcatataataagtagatggatttgctgacccggtgaggtatatatgtat
tcctgaacatgatcagtaaacgagtcgatctggccttatccgtatgaacgtcgagatctcgggaaatacaaaagctagaaggttgagattaagtatgcagattctagaagaagacgcagcgcaagttt
gcgactacgctgaatctactgctaaaaactgccacgcccacacttcttaagaatttgatttattttcacaagctgaggaacggtagggtcgaggaactcgactacaacgttctgccttgtttatttcttaaca
aaaacttagtagccgtttgggttggaaaccacctgaccttaggtctggtagcagttatttaatttattttttttatttttatacaacttgctcgctgtttgttccccctagccctgaaacacaagctgtcaaacggt
ggaggtgataagtctaatgaatgcgataagctttatttcaattcgcaattttcgtgtggcattttggcaaaaaaaaaaactcgtcggacatacatgttgccacaaacataaagtgaatacataatgttgggt
gaacgactcatacacgattgtggcaaatcaaattcttttaacacgggacggggaaaggcgagtgaagatattttagcatatatttagcacatctgttaaatccatttttttactctccgttttcggccagatat
ggttagaaaagaaaaaaattagtacataccccatatataataagaaaaaaagagagagtcagcagaagtacggggagcttaagtgtagcaatcagaacatcacaaatagtaaataaattaataataa
taataatcatatccaaaaatatttttattcctaacctatcgcattgttacatcgagggtgaaattcaaaatagacaaaaagttggggaataaaatgtgaaaaaagtggtaaaatgtttaatagtgtgggcgtt
actgttttgtcggtgtgaggtgcgtggccaccaaagtgtttttggtataacgatagaaattggtaagacaaacaatattgcgaagaaaacccgaagcatttttaaaaagtgcgaacgtggcagttttaag
ggtttgtgggcgtggcaataatttttggcaattcgataaaaatgtacaggaccaaatatatgaagaaatataaaatatttttcaaaatgacagccagcaaccatacatatatataaataaatgtcggagacc
cttccttctacctgtaacatacttttccacgaatctagtattggttgatatatatattatgctgtgtataagaccaaaatcagtgtacatttccattggattcaccaaccggatggttccggatggtaatgcaaaa
tattcatctaagaaacgaaaacacctagaattaaacctgaactgatatgacttatgcacatatcagtgaggtgggcagttcaaagcaatcacgatgctccaagttattatcgcagtgcagtgaaaaattc
acagtcaccgtcgccaattgccaataaagatcggccattatacaacagaaccgcgttgaagacgatcgacgaggtcgtgggtcttatcttatcaccacctgaattgaggcatgcctccagaatgacga
gggcatccgaagataatgtggcccgctattttcggccgggactggacctatgcgacgacctatgctgatgacgggagtctgccgctgatatggtgcaatgcaaggctccagtcgggggtataaaag
acccagtttcggtgcagtcaagacaacagactttaggtgttggtcgttgagcgaaccaaagccggagcagttgaggaaccaaagaatagcagcgagaggaccaagg

SEQ ID NO:10 (5' to 3')

tgtagggacccaaatccaattgtagtagttaccttgattatggttggcttgtccttcgatagttttgccttttccaaagcgctagaaatggattccatatcgtcgtctcctttatcgacttccatgacttcccaac
catatgcctcgtatcgcttcaaaacatcttcgtcgaacgagtacgaggttttaccgtcaatggaaatgctattactgtcataaaacgtaatcaagttacccaattgcagatgtcccgctaaggaagaggtc
tccgaagaaacacccctcttgtaagcaaccatcccctacaatagcaaacgtatatgagtcggaaatgggaaagccatcctcgttataagtggcggcaaagttggcctgcgctattgccataccaacag
catttgagatacccctggcctagcggaccggaagtgatttccactcccgctgagtggaattctggatgacccggtgtccttgagtttacttgtctaaattgtctcaagtcctcgatagagtaatcgtatccta
atagatggagcattgagtacagaagagcgcatgagtgaccgttcgacagaacaaacctgtctctattgatccaatgttcattgttagggttacagcgcagttgcttgaaaattacatgggcaactggtgc
caatcctagtggtgcacctgggtggccagattgtgcgctttccacctggtcaacggaaagtaatcttaaagtggaaaccgcaagtttatcaatgtcggagaactgtgccatttttttgttctttttttgattag
taaggtataatcgtctacgtagaggttacaaatcgaagactacagtaagaggggacaagccaattgaatatacgactgaaataaatggaataattctgcattattacactcgtttatatatccaaacaggt
gatctggtattctcttgacaacgaatgaagctccctatattcgacactccttattcaggactcctcccaacaaggagaagtaggtgttccttgagctacccttaaagctggggagatgagcttgcccttc
ctgtcatcgccattatgacgagaaaagtaaaacatgtagaataaggtccaccccaaacatgtccgagcaatgacgttatatatcgtgttccctgttcaaagcatggcatatgtgccattaaaggcgaatttt
tgtccctagcaaaggagagacagcgagccaccattaagaagtgacttgaaagcaagcgaaaatagctacacatatatatcaatatattgacctataaacccaaaatgtgaaagaaatttgataggtca
agatcaatgtaaacaattactttgttatgtagagttttttagctacctatattccaccataacatcaatcatgcggttgctggtgtatttaccaataatgtttaatgtatatatatatatatatatatggggccgtat
acttacatatagtagatgtcaagcgtaggcgcttcccctgccggctgtgagggcgccataaccaaggtatctatagaccgccaatcagcaaactacctccgtacattcatgttgcacccacacatttata
cacccagaccgcgacaaattacccataaggttgtttgtgacggcgtcgtacaagagaacgtgggaacttttaggctcaccaaaaaagaaagaaaaaatacgagttgctgacagaagcctcaagaa
aaaaaaaattcttcttcgactatgctggaggcagagatgatcgagccggtagttaactatatatagctaaattggttccatcaccttctttttctggtgtcgctccttctagtgctatttctggctttcctatttttt
ttttccatttttcttttctctctttctaatatataaattctcttgcatttttctatttttctctctatctattctacttgtttattccctcaaggttttttttaaggagtacttgttttagaatatacggtcaacgaactatataatta
actaaa

SEQ ID NO:11 (5' to 3')

tctgctattattgatgctttgaagacctccagacaaatttttcacagaatgtactcttacgttgtttaccgtattgctttgtctctacatttggaaatcttcttgggtctatggattgctattttggataactctttgga
cattgatttgattgttttcatcgctattttcgctgatgttgctactttggctattgcttacgataatgctccttactctccaaagcccgttaaatggaacctaccaagattatggggtatgtctattattttgggcat

agttttagctataggttcttggattaccttgactactatgttcttaccaaagggtggtattatccaaaacttcggtgctatgaacggtattatgttcttgcaaatttccttgactgaaaactggttgattttcattac
cagagctgctggtccattctggtcttctatcccatcctggcaattggctggtgccgtcttcgctgtcgacatcatcgctaccatgtttaccttattcggttggtggtctgaaaactggactgatattgttactgt
cgtccgtgtctggatctggtctatcggtatcttctgtgttttgggtggtttctactacgaaatgtccacttctgaagccttgacagattgatgaacggtaagccaatgaaggaaaagaagtctaccagaag
tgtcgaagacttcatggctgctatgcaaagagtctctactcaacacgaaaaggaaacctaatcctgttgaagtagcatttaatcataatttttgtcacatttaatcaacttgatttttctggtttaatttttctaat
tttaattttaatttttttatcaatgggaactgatacactaaaaagaattaggagccaacaagaataagccgcttatttcctactagagtttgcttaaaatttcatctcgaattgtcattctaatattttatccacaca
cacaccttaaaatttttagattaaatggcatcaactcttagcttcacacacacacacacaccgaagctggttgttttatttgatttgatatataattggtttctctggatggtacttttctttcttggttatttcctatttt
aaaatatgaaacgcacacaagtcataattattctaatagagcacaattcacaacacgcacatttcaacttttaatatttttttagaaacactttatttagtctaattcttaatttttaatatatataatgcacacacac
taatttattcattaatttttattgagtaggatttgaaaatatttggtatctttgcaagatgtttgtatagagggacaaagaatcgtctttattatggtcaaggctttacgtcataatagttcctgcccagctcttctat
aatactttaaagatctcttctcgtttgctccatttggaagtctcgcttacgtttatgcgcccatacagacactcaagatacacacttacatgaacgtatacaaatttactaacactacttgaaaatatgaacca
cagtacatcatattaagacgtagtattcgatgattgaaggccgcctccgcgaaatacctttactgattttgccggttaatcgcatcgaaatttcttcatcacaagaaagcaaacaaatcgccaggccattct
acaagtttccttttcttatgaagatgtaaaagctactaaggcgtcattactctagatgactcagtttagtctgaccttctatagtatactaccctggcgctatgatgatgagcggttcttttattgcggaaacga
aaattccgggaccggcgaaatttgcccggttttgtccgtaaccggcttcatgagtcggcttcaatagtagttgaatacttatttaaacagcagaacttaactcactcatcacgctgtttccgctgaattttctc
aaaatatctaagcagtcaacaaatataaagaatattgaaattgacagtttttgtcgctatcgatttttattatttgctgttttaaatc

SEQ ID NO:12 (5' to 3')

ccaaatcattcttattcggtttccagacggtaacaatacccctcgcccatcccacacaaaagggtatctgctacttcgggatcgacgaaacaaccacaaagaacttcgtcctcctgatcatcgctgatcaa
aattttaccatcctcgtttccagctacgttcggtttggcatctttgtcgcgaacgtcaaaataagctaacgttgtctggcccaaagaaatgaatttatatgcagatcttttatcaaagtggaaaatatcgttgat
agagtcgccaaaatgtatcgaacgaatggaatttgataatgccaagttttccgagtttattacgtgtatattaccggattcatcgcctattaaaatgaatgggtgagtttgagaggcgcataatttcgtaaat
ttatcatttttttttctcttcagaattaaacagtgagcttaagtttacctttttgacgactttgccggtcatagtattggccttttttaaaacattatccgatccaacagaaaaaatattgtcaccttttagaatcaaagc
acatggcacggacactacctttatgtctttagtcttccaaagtgtttttacgcccaagtcttcatcttttcctgtttgcttttgttgctgttgttcttcaatatcaacaaatttcaaatctccagtttctaggtctatat
ctaatctaatccaagggcatacacctttctttgcatccttgcctgtagttgcagtgtcaattctacgtctacgatctaggtgcgattgcaacttagcggggtcataacgatggcacacaatatgtcctgtacc
aaagccagttattataatgggcagttcaggatgtaaaagagactggaaaatgggagcttttaatgatagtaattctagaatgggcaggtttgttgaatcgacaacatctgttttttttttgctctttgccatag
ctgatgcgtggattgtttctaatttcccagctgcttcctcttccaattgtggcgatgatgccatgatttctatgttaaaatttttctaaccatgaaatttttttttctagcgagaaaaaaaatcagaaaaattactat
tagtgagtattggagacattgtcaatgggagatgttctctttataatatcttcaacaggttctttcaactctggaaattcatccacaatcttgtcagcaagtgaatctcttaattgcttcaatccatgcatcttgcc
tctttgatattggttggatcttcttatggcttccacgaactctcttgtgtaaatatctggatttctaccgtcctcaatgtattgaacaacttccaagggaatgtccaccttagacaagctggattgaggatcgttg
cttctcacgttcagcttgtacaagcgatccacatttctttgcaagttggtgatcattcccttggtggcttctggagtaccaggaaaatcatatatcgagacacctaattcaacgaaggactcaataatcgaa
gccacttggtcttgagtagtggccagttcttgctgcaattgttcattgttagtgctgtttccattcatcttatcggtttatttttctatatatttgcctctttctcaaacaggagttagtagttaaaagtacgaagttct
tgttctttaatgcgcgctgacaaaagaattggataaaagagaatggtgggggacaagaaggaaatttgtcctagtttaacatgaatggcatcttgttaccgggtggacatcacctattgattctaaatat
ctttacggtttatcatactgttctttattccgtcgttattcttttttatttttatcatcatttcacgtggctagtaaaagaaaagccacaacatgactcagcaaatctcgacaaagtaaaagctcatagagatagtat
tatattgatataaaaaaagtatactgtactgtttgtaaccttttcaatgctttaagatcaaaactaaggccagcaaaggtatcaacccatagcaactcataaa

SEQ ID NO:13 (5' to 3')

gaaaccattaaatcatatttaataaattgttgcgacatgcaagaagttcgcggatggtcatgcgtatttaagaatagtcaagtaacaatttgcttattcgttgatgatatgatattattcagcaaagacttaaat
gcaaataagaaaatcataacaacactcaagaaacaatacgatacaaagataataaatctgggtgaaagtgataacgaaattcagtacgacatacttggattagagatcaaatatcaaagaagcaagta
catgaaattaggtatggaaaaatccttgacagaaaaattacccaaactaaacgtaccttgaacccaaaaggaaagaaactagagctccaggtcaaccaggtcattatatagaccaggatgaactag
aaatagatgaagatgaatacaaagagaaagtacatgaaatgcaaaagttgattggtctagcttcatatgttggatataaatttagatttgacttactatactacatcaacacattgctcaaccatatactattc
ccctctaggcaagttttagacatgacatatgagttaatacaattcatgtgggacactagagataaacaattaatatggcacaaaaacaaacctaccaagccagataataaactagtcgcaataagcgat
gcttcatatggtaaccaaccatattacaagtcacaaattggtaacatttcctactcaacggaaaagtgattggaggaaagtcgacaaaggcttcgttaacatgcacttcaactacagaagcagaaatac
acgcggtcagtgaagctattccgctattgaataacctcagtcaccttgtgcaagaacttaacaagaaaccaattattaaaggcttacttactgatagtagatcaacgatcagtataattaagtctacaaatg
aagagaaatttagaaacagatttttggcacaaaggcaatgagacttagagatgaagtatcaggtaataatttatacgtatactacatcgagaccaagaagaacattgctgatgtgatgacaaaacctctt
ccgataaaaacatttaaactattaactaacaaatggattcattagatctattacattatgggtggtatgttggaataaaaatcaactatcatctactaactagtatttacgttactagtatattatcatatacggtg
ttagaagatgacgcaaatgatgagaaatagtcatctaaattagtggaagctgaaacgcaaggattgataatgtaataggatcaatgaatattaacatatataaaatgatgataataatatttatagaattgtgt
agaattgcagattccctttttatggattcctaaatcctcgaggagaacttctagtatatctacatacctaatattattgccttattaaaaatggaatcccaacaattacatcaaaatccacattctcttcaaaatca
attgtcctgtacttccttgttcatgtgttcaaaaacgttatatttataggataattatactctatttctcaacaagtaattggttgtttggccgagcggtctaaggcgcctgattcaagaaatatcttgaccgca
gttaactgtgggaatactcaggtatcgtaagatgcaagagttcgaatctcttagcaaccattatttttttcctcaacataacgagaacacacaggggcgctatcgcacagaatcaaattcgatgactgga
aatttttgttaatttcagaggtcgcctgacgcatatacctttttcaactgaaaaattgggagaaaaaggaaaggtgagagccgcggaaccggcttttcatatagaatagagaagcgttcatgactaaatg
cttgcatcacaatacttgaagttgacaatattatttaaggacctattgtttttccaataggtggttagcaatcgtcttactttctaacttttcttacctttacatttcagcaatatatatatatatttcaaggatata
ccattctaa

SEQ ID NO:14 (5' to 3')

cactaccaccactacggttgtccatgacgtatcctgcgatttttgaattaatgattcaatagttgacatttgctcgtcattggggttcgactgagctgcggatgtcaacttcgcaacagcttctgcatggttt
ccttgagaaaaatgagactcagcctctgagattaacttatccgtatccatttcagatctttgctatacgtttgtatcgctatatgtacgttcttttaatgaacttttctccttttctttatcgtgtagcttgcttgggtatc
ttttaatgagttgcggacagtgagatttttcagaagggcaattggccaagacaccaaaaacgtttggacgagacaggcatcaaaggacaaggtaaaaggcgttgagctgtggctggctgtgtatgcg
tttgaaataccatggatagatatcaaagaaagatgatggatgtttcatacaaatcccaaatttggggcgcggacaactgaaatacgtgggtccagtggacacgaaagctggaatgtttgctggtgtagact
tacttgccaacattggtaagaacgatggatcattcatggggaagaagtattttcaaacagagtatcctcaaagtggactatttatccagttgcaaaaagtcgcatcattgatcgagaaggcatcgatatcg
caaacctcgagaagaacgacgatggaaccgctatcaatacccaaaaacagatctattgtgaggctcactaaccagttctctcccatggatgatcctaaatccccacacccatgagaagtttccggat
caccagtcggcacagcggtaatcaacagtcgatggaccaggaggcatcggatcaccatcaacagcaagaatttggttacgataacagagaagacagaatggaggtcgactctatcctgtcatcag
acagaaaggctaatcacaacaccaccagcgattggaaaccggacaatggccacatgaatgacctcaatagcagcgaagttacaattgaattacgagaagcccaattgaccatcgaaaagctacaa
aggaaacaactacactacaaaaggctactcgatgaccaaagaatggtcctcgaagaagtgcaaccgacttttgataggtatgaagccacaatacaagaaagagagaaagagatagaccatctcaa
gcaacaattggagctcgaacgcagacagcaagccaaacaaaagcagttttttgacgctgagaatgaacagctacttgctgtcgtaagccaactacacgaagagatcaaagaaaacgaagagagaa
atctttctcataatcaacccactggtgccaacgaagatgtcgaactcctgaaaaaacagctggaacaattacgcaacatagaagaccaatttgagttacacaagacaaagtgggctaaagaacgcga
acaattgaaaatgcataacgattcgctcagtaaagaatacccaaaatttgagcaaggaactatttttgacaaaaccacaagattcctcatcggaagaggtggcatcctttaacgaaaaaacttgaagaggc
taatgaaaaaatcaaacagttggaacaggctcaagcacaaacagccgtggaatcgttgccaatttttcgacccccctgcaccagtcgataccacggcaggaagacaacagtggtgtgagcattgcga
tacgatgggtcataatacagcagaatgcccccatcacaatcctgacaaccagcagttcttctaggcagtcgaactgactctaatagtgactccggtaaattagttaattaattgctaaacccatgcacagt
gactcacgttttttatcagtcattcgatatagaaggtaagaaaaggatatgactatgaacagtagtatactgtgtatataatagatatggaacgttatattcacctccgatgtgtgttgtacatacatataaaaat
atcatagcacaactgcgctgtgtaatagtaatacaatagtttacaaaatttttttctgaatac

SEQ ID NO:15 (5' to 3')

acaatgaggaagaacatgccgttttacaagaattaaatagtttaacccaaagaattaatgaactaggcatggaaagtataaattcaaactccgattcggacagaataaacgggtcatattcacaagtgg
attttggtaacaataacgacgaggacgatatgaacctgttcgacccagattttatggcacaagaccaattgcgtgctgaagaaagagactacaacaaggatgatagaacacccttagctaaggtccct
gcggcctttcaatcaactggattgggcataacccccgatgacgatatcgagagacaatacataacggaacacagatcacgacatgaagtgccaaagcggtctcccgagaaaccctccaaccccgct
ggaaatataggtaacccatacgcgaaacctggcacaaggttgaataccactcacacccacagcaaaactgatcgtagcattacccctcagaggggccagccagtcccatcaggccagcagatttcct
cctacgtgcagccagcaaacattaatagtcctaacaaaatgtatggtgcaaacaactcggcaatgggttcgcccaggaatccaaagacgagagcgccaccaggtccatacaatcagggatggaat
aaccgccccctcgccttcaaatatttaccaacgtcctcatccctcagatacacaaccacaagcatatcatctccccggaaacccatactcaacggggaacaggccaaacatgcaagcgcaatatcacc
cgcagcaggtgcccatgcctatcctgcagcagcccaatcgcccgtaccaaccttatgcgatgaatacgcacatgggctctcctggcggatatgctggggcagcaccaccatttcagccagctaacg
tcaactacaatactaggcctcagcagccatggcctacacctaactcaccatccgcacactaccgtccgccccctaacctgaaccagcctcaaaacggtagtgctggttactatcgtccgccggcacc
acaattgcaaaactcccaagcccgtccacaaaagaaggacggattctcacagttcatgccatctgcaactacgaagaacccatatgcccagtaactcgaccgactggttgtaattttacaaaaagaga
gacaatttaagaaaagaaacaagcgccaggcttccgtatcccagttttcatctcactttctgggcacgattgtaataatacttcatgataaataactaaactatataagtagtgtctcatccgtaaatatacatt
tagacagattcttgtattttctccgggcaatttttaacttttttctgttagggcacatgacacttgcctattatggacagccagtaaagatgtgccatatattgccccctttacgctctctgccagtattagtgg
gaaaaaaaaaactgaaaaaaaaaaaatcgcagactactaataatcacgtgatatttcttttcactctcttcataaagttgctaaaaacacacaatcgaatgagcctctgagcagtataaattgtacttcaaa
gcactagtcatgaaaaacgcttacattagttcagtttgtcaaggttatgctattacttgtacttatttcttgctattgttagtggctcccccacattgacgtattttcacgtgatgcgcctcactgcggaaggcgc
cacacattgcctgcaaaaaattgtggatgcactcatttgatagtaaactaagtcatgttaatcgtttggatttggcacacacccacaaatatacacattacatatatatatatttcaaaatacagctgcgtcc
aatagatgagcttccgcttcgttgtacaacctacctgctatcttgttcacggatatttcttgcttttaataaacaaaagtaactctagaacagtcaagtcttcgataatttttttagtcacagggtccgtctaaag
tttctctttatttggaataatagaaaagaaagaaaaaaacgtagtataaaagga

SEQ ID NO:16 (5' to 3')

aaggatggcaaatacccaatcggaggaactcgaacacttcagtatctgtgtcttctagtgagtctttagcggaagttattcagccatcttccttcaaaagtgggagtagttcattgcattatctatcgtcttct
atctcaagccaacctggttcgtacggttcttggttcaacaaaaggccaacaatttctcagttctttcaaccaagcccttctttaaaacacaacgagtcgtgggagaggctgcaaacaactgctggaaata
tgcaaaggacttcaagttcgtcttctttgcagcaagcaacctccaggttatcactaaccactccgcaacaatcaccgtctatcagcgaatatgatgagtatccttggatgggcacacctggctctcctaat
gttggagatgtgtctcacgcacccccattggttaagaatatatcatataaatttccactaaagaacgttgagttgaaaagagattgccaaaggatctctcaggatgatcttttggatgaggcttttgaaaga
atatgtcagccctctttggctgaccttaattccacttacgaaatttttccaggtaactcttcttatgcggatattttgactactgattctgatattgatgatggcttgatgaataaacctctggaactattgccgaa
atatacaatgtatttaacccatttttaacaatttttccagttgcaagcatgtcctgctggtcaagaatcagagagcagaataacaaattctatgaagattgacctgttaaaggcggattacacaagaagtcta
ttagtatcgttacgttcaagggacattagggatgtcgcattgaaaagagagtttactggcaataacaacaataacagcaaccagaatatctatgatgagaatttttgtcggaaaaaggaagtacgtgttga
aacagaagaccagaaaaatctttttcctgtggcaagattggcaagctaagtactagtttggaaaactgcgttaattttgttgaaaatagtataaagagtgcaatgatgttatatgatgataatggaatagata
gtgagcttcgcgattcagaagctttacggatttttttcatctcttgttcattattgtaatgcaggttaatgttttctccttctttacatgtttaatatattccaagttacctaagaggtgtacgatatttttttctttatata
tatgattttctctattcatttttagttttttttgatacataagcgaatcgcacattgcgcaacttcaatttgttgattcgccaaagtattcttaccataaaacaaccattcgttgctttacccttcgtaatcatttacc
gtgataaccataatcagaaacttattatttcagcctagtagaccggccaagcaggccttgtaatgtttctcttgattgcttgaatcttttaagcagccaaatctttccaaaaaaaatgcaattatcagaacaaaa
ctatttaaggtgacttctccgtatttacaccaccagaagcgttctggctccccttttctctaaacgttaaacatttttacaattgaaatgttaccaatcctatattattgtaccacattgccagatttatgaactctg
ggtatgggtgctaatttttcgttagaagcgctggtacaattttctctgtcattgtgacactaattaggaaacttctcgactatcaatgtgtaaatgaaggaataatggcggaaacttttgaaacttttgtcaataatt
gcatcattggatgcgtttcatttggccgttatcacggagaggcagagttctctccacaatttgggcagaagtcttttgaaaagacatatatatatatatatatgtatatgagtggatgcttaaggtaagaata
atttctgaattcccaagtattcattttgtgcagtattcacatattctatttttattgcttttaacttagaggcaattaaatttgtgtaggaaaggcaaaatactatcaaaattttcc

SEQ ID NO:17 (5' to 3')

ttgccttcaagatctactttcctaagaagatcattattacaaacacaactgcactcaaagatgactgctcatactaatatcaaacagcacaaacactgtcatgaggaccatcctatcagaagatcggactc
tgccgtgtcaattgtacatttgaaacgtgcgcccttcaaggttacagtgattggttctggtaactgggggaccaccatcgccaaagtcattgcggaaaacacagaattgcattcccatatcttcgagcca
gaggtgagaatgtgggtttttgatgaaaagatcggcgacgaaaatctgacggatatcataaatacaagacaccagaacgttaaatatctacccaatattgacctgccccataatctagtggccgatcct
gatcttttacactccatcaagggtgctgacatccttgttttcaacatccctcatcaatttttaccaaacatagtcaaacaattgcaaggccacgtggccccctcatgtaagggccatctcgtgtctaaaaggg
ttcgagttgggctccaagggtgtgcaattgctatcctcctatgttactgatgagttaggaatccaatgtggcgcactatctggtgcaaacttggcaccggaagtggccaaggagcattggtccgaaacc
accgtggcttaccaactaccaaaggattatcaaggtgatggcaaggatgtagatcataagattttgaaattgctgttccacagaccttacttccacgtcaatgtcatcgatgatgttgctggtatatccattg
ccggtgccttgaagaacgtcgtggcacttgcatgtggtttcgtagaaggtatgggatggggtaacaatgcctccgcagccattcaaaggctgggtttaggtgaaattatcaagttcggtagaatgttttt
cccagaatccaaagtcgagacctactatcaagaatccgctggtgttgcagatctgatcaccacctgctcaggcggtagaaacgtcaaggttgccacatacatggccaagaccggtaagtcagccttg
gaagcagaaaaggaattgcttaacggtcaatccgcccaagggataatcacatgcagagaagttcacgagtggctacaaacatgtgagttgacccaagaattcccattattcgaggcagtctaccaga
tagtctacaacaacgtccgcatggaagacctaccggagatgattgaagagctagacatcgatgacgaatagacactctccccccccctccccctctgatctttcctgttgcctcttttccccccaaccaa
tttatcattatacacaagttctacaactactactagtaacattactacagttattataatttttctattctctttttctttaagaatctatcattaacgttaatttctatatatacataactaccattatacacgctattatcg
tttacatatcacatcaccgttaatgaaagatacgacaccctgtacactaacacaattaaataatcgccataacctttttctgttatctatagcccttaaagctgtttcttcgagcttttttcactgcagtaattctcca
catgggcccagccactgagataagagcgctatgttagtcactactgacggctctccagtcatttatgtgatttttttagtgactcatgtcgcatttggcccgttttttttccgctgtcgcaacctatttccattaac
ggtgccgtatggaagagtcatttaaaggcaggagagagagattactcatcttcattggatcagattgatgactgcgtacggcagatagtgtaatctgagcagttgcgagacccagactggcactgtct
caatagtatattaatgggcatacattcgtactcccttgttcttgcccacagttctctctctctttacttcttgtatcttgtctccccattgtgcagcgataaggaacattgttctaatatacacggatacaaaagaa
atacacataattgcataaaatac

SEQ ID NO:18 (5' to 3')

ttttgtaagaaattattcaccgcatcttcatctggcaaacgaatgggagactttgaggaacccaatccatttctgaataacggagatttagaaatgtaaaaggtagcaaatgtaaaaagtgccaggacca
tcacagcagtcaatgccaacaccaatttcccttgccatgacactgttggatcttttgaaggagatttgtaacctggaatctcactataatgaacacattcaccggattcacacttcaaagtaatataagggt
caccaaacacggtcaatatcaaatcattcatagaaggctcactgaatttacattgccttgtttctaaatcacagctgaaatctcctggcccttttattgtctctgtcaggaaatccgagatatctatagacccc
ttagcaccacacaacacagtgtcgggaacgcatttgcattgaacgtcattacacttataatgggaggtattctgttccaagtcgtattcaaaggcacaatcacttaagccacaatagaagctttctaactg
atctatccaaaactgaaaattacattcttgattaggtttatcacaggcaaatgtaatttgtggtattttgccgttcaaaatctgtagaattttctcattggtcacattacaacctgaaaatactttatctacaatcat
accattcttataacatgtcccctaatactaggatcaggcatgaacgcatcacagacaaaatcttcttgacaaacgtcacaattgatccctccccatccgttatcacaatgacaggtgtcattttgtgctctt
atgggacgatccttattaccgctttcatccggtgatagaccgccacagaggggcagagagcaatcatcacctgcaaacccttctatacactcacatctaccagtgtacgaattgcattcagaaaactgtt
tgcattcaaaaatagtagctagcatacaattaaaacatggcgggcatgtatcattgcccttatcttgtgcagttagacgcgaatttttcgaagaagtaccttcaaagaatggggtcttatcttgttttgcaagtac
cactgagcaggataataatagaaatgataatatactatagtagagataacgtcgatgacttcccatactgtaattgctttagttgtgtatttttagtgtgcaagtttctgtaaatcgattaatttttttttcttcctc
tttttattaaccttaatttttattttagattcctgacttcaactcaagacgcacagatattataacatctgcataataggcatttgcaagaattactcgtgagtaaggaaagagtgaggaactatcgcatacctg
catttaaagatgccgatttgggcgcgaatccttattttggcttcaccctcatactattatcagggccagaaaaaggaagtgtttcccctccttcttgaattgatgttacccctcataaagcacgtggcctcttat
cgagaaagaaattaccgtcgctcgtgatttgtttgcaaaaagaacaaaactgaaaaaacccagacacgctcgacttcctgtcttcctattgattgcagcttccaatttcgtcacacaacaaggtcctagc
gacggctcacaggttttgtaacaagcaatcgaaggttctggaatggcgggaaagggtttagtaccacatgctatgatgcccactgtgatctccagagcaaagttcgttcgatcgtactgttactctctct

ctttcaaacagaattgtccgaatcgtgtgacaacaacagcctgttctcacacactcttttcttctaaccaaggggggtggtttagtttagtagaacctcgtgaaacttacatttacatatatataaacttgcataa
attggtcaatgcaagaaatacatatttggtctttctaattcgtagtttttcaagttcttagatgctttctttttctctttttttacagatcatcaaggaagtaattatctactttttacaacaaatataaaaca

SEQ ID NO:19 (5' to 3')

aaacaaatggcaaaaataacgggcttcaccattgttcctgtatggtgtattagaacatagctgaaaatacttctgcctcaaaaaagtgttaaaaaaaaagaggcattatatagaggtaaagcctacaggc
gcaagataacacatcaccgctctccccctctcatgaaaagtcatcgctaaagaggaacactgaaggttcccgtaggttgtctttggcacaaggtagtacatggtaaaaactcaggatggaataattca
aattcaccaatttcaacgtcccttgtttaaaaagaaaagaatttttctctttaaggtagcactaatgcattatcgatgatgtaaccattcacacaggttatttagcttttgatccttgaaccattaattaacccaga
aatagaaattacccaagtggggctctccaacacaatgagaggaaaggtgacttttaaggggccagaccctgttaaaaacctttgatggctatgtaataatagtaaattaagtgcaaacatgtaagaa
agattctcggtaacgaccatacaaatattgggcgtgtggcgtagtcggtagcgcgctcccttagcatgggagaggtctccggttcgattccggactcgtccaaattatttttttacttttccgcggtgccga
gatgcagacgtggccaactgtgtctgccgtcgcaaaatgatttgaattttgcgtcgcgcacgtttctcacgtacataataagtatttttcatacagttctagcaagacgaggtggtcaaaatagaagcgtcc
tatgttttacagtacaagacagtccatactgaaatgacaacgtacttgacttttcagtattttcttttttctcacagtctggttatttttgaaagcgcacgaaatatatgtaggcaagcatttttctgagtctgctgac
ctctaaaattaatgctattgtgcaccttagtaacccaaggcaggacagttaccttgcgtggtgttactatggccggaagcccgaaagagttatcgttactccgattatttttgtacagctgatgggaccttgc
cgtcttcatttttttttttttttcacctatagagccgggcagagctgcccggcttaactaagggccggaaaaaaaacggaaaaaagaaagccaagcgtgtagacgtagtataacagtatatctgacacgc
acgtgatgaccacgtaatcgcatcgcccctcacctctcacctctcaccgctgactcagcttcactaaaaaggaaaatatatactctttcccaggcaaggtgacagcggtccccgtctcctccacaaag
gcctctcctggggtttgagcaagtctaagtttacgtagcataaaaattctcggattgcgtcaaataataaaaaaagtaaccccacttctacttctacatcggaaaaacattccattcacatatcgtctttggc
ctatcttgttttgtcctcggtagatcaggtcagtacaaacgcaacacgaaagaacaaaaaaagaagaaaacagaaggccaagacagggtcaatgagactgttgtcctcctactgtccctatgtctctg
gccgatcacgcgccattgtccctcagaaacaaatcaaacacccacaccccgggcacccaaagtccccacccacaccaccaatacgtaaacggggcgccccctgcaggcccctcctgcgcgcggc
ctcccgccttgcttctctccccttccttttctttttccagttttccctatttgtccctttttccgcacaacaagtatcagaatgggttcatcaaatctatccaacctaattcgcacgtagactggcttggtattggc
agtttcgtagttatatatatactaccatgagtgaaactgttacgttaccttaaattctttctccctttaattttctttttatcttactctcctacataagacatcaagaaacaattgtatattgtacaccccccccctcc
acaaacacaaatattgataatataaag

SEQ ID NO:20 (5' to 3')

ggatgagaaacgagtgcggtttcgagagtagatattcaacccacccgaagtagccttcaggaactggttccgttctctcttcctccggaatagtctgaatgtccttaagagaccgtggctcgtatactctt
ctattcttgggccgcaatagcaaaaagagccagacaaacacgacggcggtaagaccgtagataatcagggttgaaatgaacgccgaagtcgaagaactgtcagccatagtacgtatgtgctataaa
tatctaacctttcgctgctttgaatatgatgtgctcaaatataacttaatataatagtataacaaaaaggagtactatttgctaaatatcgtagacgtagtagacatagtaaatacaataaaggatagataacc
aagaacccacatcaagcgaatacatacatatatatatactcgatgtatacatgtttctaagcacttgcgcacatacgtatttaaagtatttcagggagattaacgtattaaaacaagaagagggttgactac
atcacgatgaggggatcgaagaaatgatggtaaatgaaatagaaatcaaggagcatgaaggcaaaagacaaatataagggtcgaacgaaaaataaagtgaaaagtgttgatatgatgtatttgg
ctttgcggcgccgaaaaaacgagtttacgcaattgcacaatcatgctgactctgtggcggacccgcgctcttgccggcccggcgataacgctgggcgtgaggctgtgcccggcggagttttttgcg
cctgcattttccaaggtttaccctgcgctaaggggcgagattggagaagcaataagaatgccggttggggttgcgatgatgacgaccacgacaactggtgtcattatttaagttgccgaaagaacctg
agtgcatttgcaacatgagtatactagaagaatgagccaagacttgcgagacgcgagtttgccggtggtgcgaacaatagagcgaccatgaccttgaaggtgagacgcgcataaccgctagagta
ctttgaagaggaaacagcaatagggttgctaccagtataaatagacaggtacatacaacactggaaatggttgtctgtttgagtacgctttcaattcatttgggtgtgcactttattatgttacaatatggaa
gggaactttacacttctcctatgcacatatattaattaaagtccaatgctagtagagaagggggtaacaccccctccgcgctcttttccgatttttttctaaaccgtggaatatttcggatatccttttgttgttt
ccgggtgtacaatatggacttcctctttttctggcaaccaaaacccatacatcgggattcctataatacctttcgttggtctccctaacatgtaggtggcggaggggagatatacaatagaacagataccaga
caagacataatgggctaaacaagactacaccaattacactgcctcattgatggtggtacataacgaactaatactgtagccctagacttgatagccatcatcatatcgaagtttcactacccttttttccattt
gccatctattgaagtaataatagggcgcatgcaacttcttttcttttttttctttctctctccccccgttgttgtctcaccatatccgcaatgacaaaaaaatgatggaagacactaaaggaaaaaattaacgac
aaagacagcaccaacagatgtcgttgttccagagctgatgaggggtatctcgaagcacacgaaacttttccttccttcattcacgcacactactctctaatgagcaacggtatacggccttccttccagt
tacttgaatttgaaataaaaaaaagtttgctgtcttgctatcaagtataaatagacctgcaattattaatctttgtttcctcgtcattgttctcgttcccttttcttccttgtttctttttctgcacaatatttcaagctata
ccaagcatacaatcaactatctcatatacaatgtctatcc

SEQ ID NO:21 (5' to 3')

ggcagtcatcaggatcgtaggagataagcaccctgacaagtaacatgccgatgaagttgtttggttcattgggcaaaaaaatcgggattctagaaaaccctgagttgaagatttttcgacagttttatc
gtctaggatggtatcggcactcattgtgaacacgttttcaatcggagtcatgatttcctcaaccctctttgccttagatccaaaacagcagagatgattgtaacttcgtcttagtcaaccgttccaccccc
atggtcctatgcaaggtgaccaaagtctttaagccggatttttgtacatcgtaccatgatcttcacccagcatatagtccaggagagtcgcgatcggatatgcgactgggtacatcagatacatcagta
caagaacaaaggggcagaagaatgccccaacttgcagcccgtatttaacacagacactctgcggaataatttcaccgaagatcacaattagaatagttgacgacactacagcctgccaaccacccc
caagacacctgtccaaaacaataggcaatgtttcgttggttataacattagaaagcagcagtgtgactagaacccaatgcttccccctagatattaggtcaagcacccgcttggccagtttcttttcagaa
ttcgagcctgaagtgctgattaccttcaggtagacttcatcttgacccatcaaccccagcgtcaatcctgcaaatacaccacccagcagcactaggatgatagagataatatagtacgtggtaacgctt
gcctcatcacctacgctatggccggaatcggcaacatccctagaattgagtacgtgtgatccggataacaacggcagtgaatatatcttcggtatcgtaaagatgtgatataagatgatgtatacccaat
gaggagcgcctgatcgtgacctagaccttagtggcaaaaacgacatatctattatagtggggagagtttcgtgcaaataacagacgcagcagcaagtaactgtgacgatatcaactctttttttattatgt
aataagcaaacaagcacgaatggggaaagcctatgtgcaatcaccaaggtcgtcccttttttcccatttgctaatttagaatttaaagaaaccaaaagaatgaagaaagaaaacaaatactagccctaa
ccctgacttcgtttctatgataatacctgctttaatgaacggtatgccctagggtatatctcactctgtacgttacaaactccggttattttatcggaacatccgagcacccgcgccttcctcaacccaggc
accgcccccaggtaaccgtgcgcgatgagctaatcctgagccatcacccaccccacccgttgatgacagcaattcgggagggcgaaaaataaaaactggagcaaggaattaccatcaccgtcac
catcaccatcatatcgccttagcctctagccatagccatcatgcaagcgtgtatcttctaagattcagtcatcatcattaccgagtttgttttccttcacatgatgaagaaggtttgagtatgctcgaaacaat
aagacgacgatggctctgccattgttatattacgcttttgcggcgaggtgccgatgggttgctgaggggaagagtgtttagcttacggacctattgccattgttattccgattaatctattgttcagcagctc
ttctctaccctgtcattctagtatttttttttttttttttggttttacttttttttcttcttgccttttttcttgttacttttttctagttttttttccttccactaagcttttttccttgatttatccttgggttcttctttctactccttag
attttttttttatatattaatttttaagtttatgtatttttggtagattcaattctctttcccttttccttttccttcgctcccctttccttatca

SEQ ID NO:22 (5' to 3')

tgacaacgagtaccaggaaatcagtgcttctgctttgaagaaggctcgtaagggctgtgatggtttgaagaaaaaggcagtcaagcaaaaggaacaggagttgaagaaacaacaaaaagaggca
gaaaatgctgccaagcaattgtctgcttgaatatcaccattaaggaggacgaatcgctaccagctgccattaagactagaatttatgactcttattccaaggtcggacaaagagttaaggtttccggttg
gatccatagattacgttctaacaagaaggttattttcgtcgtcctcagagacggatctggtttcattcaatgtgtcttgtccggtgatttggcattggctcaacaaacttttggacctgacttttggaatccaccg
ttactctgtacggtaccatagtcaaattgcctgagggtaaaaccgctccaggtggtgttgaattgaatgtcgactattacgaagttgtaggtttggccccccggtggtgaagactcctttacaaacaaaatc
gcagagggctcagaccctttcttactgttggaccaacgtcatttggccttgagaggagatgccttgtctgcagtcatgaaagtccgtgctgctctactgaaaagcgttagacgtgtttatgatgaagaac
atttgacagaagttaccccaccatgtatggtgcaaactcaagtcgaaggtggttccacttttgttcaagatgaactattacggcgaggaagcttacttgacccaaagttcccaattatatttagaaacctgtt
tggcctccctaggtgatgtttataccatccaagaatctttcagagctgaaaagtcccacacaagaagacatttgtccgaatatacccatatcgaagctgaattggccttcttgactttcgacgatctattac
aacatattgaaactttgatcgtcaaatccgtgcaatacgttttggaagacccaattgctggcccactcgtaaaacaattgaatccaaactttaaggctccaaaggctccattcatgagattacagtacaag
gatgccattacctggttgaacgaacacgacatcaagaacgaagagggcgaagactttaaatttggtgacgatattgcagaagctgctgaaagaaagatgaccgataccatcggcgtcccaatctttttt
gacgagattcccagtagaaatcaagtctttctacatgaagcgttgttctgacgacccccgcgtcactgaatccgtcgacgttttgatgccaaacgttggtgaaatcactggtggttctatgagaatcgac
gacatggacgaactaatggcagggtttaagcgtgagggtattgataccgacgcctactactggttcattgaccaaagaaaatacggtacttgcccacatggtggttacggtatcggtaccgaacgtatt

ttagcctggttgtgtgacagattcactgtcagagactgttccttgtatccacgtttcagcggtagatgtaagccatgatctttagttactgaagagtacgtgagcgctcacatatatacaaatatttataccga
ttaatatttacgttcctccctctctctaattattcattgatttattcaagaattagcgttataacaataaatggttggcgcaggcaattaatttttctttactcttccaaaccctctgttaacgacaatcaaataacct
gatctgccaaggctccatcatatctggcctagaacagttttttttttcgattatttgttcgttcttgtggtggttactcattggcagaatcccgaaaatcatgattagtagatgaatgactcacttttttggataag
ctggcgcaaattgaaacatgtgaaaaaaaaaaaaaaggattataaaaggtcagcgaagcacagaactctgagataagactacctttctttagctaggggagaatattcgcaattgaagagctcaaaa
gcaggtaactatataacaagactaaggcaaac

SEQ ID NO:23 (5' to 3')

tcctaaggacatattccgttcgtacttgagttattggatctatgaaatcgctcgctatacaccagtcatgattttgtccctggtaataggggttttggttttattaattatattttttaatgacaacgaagcttgtgtt
ttcaattctgcaatatttgcttttacttctcttgtaggtttgttaataatattaagtgatggtaatccaaagctagtcagtcgtcgaaattttaggaccgagcttttagtggatgtcatcacacgtaaaccggcgg
tagaagggaaagaatggaggatcatcacatacaacatgaaccaatatttgtttaatcatgggcaatggcatactccgtattacttttacagcgatgaggattgctaccgttattttctacgccttgttgagg
gagtaacccccaagaagcaaacagccacgtcaattggcaattctccggtcaccgctaagcctgaagatgccatcgagtcagcttctcctagttccagactgaattatcaaaacttttttgctcaaggcag
cggagatcgaacgacaagctcaggaaaattactggcgaaggcggcatcccaatatcgatgcgcttcttaaaaagacggaatagcttagagacactaccatacgtaaagcgaacataaactagagta
tgatatataatcagcactaactggccggaaaacggccgaaggaagcctcgaaaagtcgattcgtgttggacccatttgctgaacaaagtggttcattgcctacctattatggtagtagtcgtgataatcg
tgtggttggttttgtcaacggtgcatttgcattttcatgacaataaaccttgcgttttcgttctcgggatattactttccctccacttctttcgcctcaatagctcctataagcattctcagggcgtatgtcggtga
tcgagatttccaagcaagcttttagtggaaatcatcgcgcgcaagccagcggtaaagggaaaagaacggaggacgattacatacaagatgaacgaataaataaattaataataaataataaaaaa
gtacagtagcattaaatattattaagtttaatgattaaaaattggttaattgtcaagaaaatctaaggtattaataaataaataatactatgacaacttgcagcgaaagcatcagccccaatgaaaattaatca
gaattgaatctgagcgtatttatttgataacggtttacgtaactgttggaataaaaatcaactatcatctactaactagtgtttacgttactagtatattatcatatacggtgttagaagatgacgcaaatgatga
gaaatagtcatcgttttcaacggaagctgaaatacaaggattgataatgtaataggatcaatgaatatcaacatataaaacgatgataataatatttatagaattgtgtagaattgcagattccctttttatgga
ttcctaaatcctcgagaagaacttctagtatatctacgtacctaatattattgccttattaaaaatggaatcccaacaattatctcaaaatcccccaattctcatcagtaacaccccaccccgtattacttttac
cgtgatgaagattggcatcgttactttctaaacgtaggacgtgcggaatgacaaaaccatcagcagtgtcacgatctctccagtcacaatggcaatcatgagtgcatagtccaaagtaaaggggcaag
gaaaagcatgattgaaaggactccccatctggactctatatgtcatcagcggctaaaaaaaagcatatagcacaacatcagcatcagcatcagcactagagtcatcggcccggcggtccgcggtcat
ccccgcggactttccgtccgcccggcgggctgtatcagcgtcaactggaacgcgcatatatatacaagacacacataacatagaagcacacccacgacaataaccacacgacaataaccacaccc
gcccacccctccttccgtatac

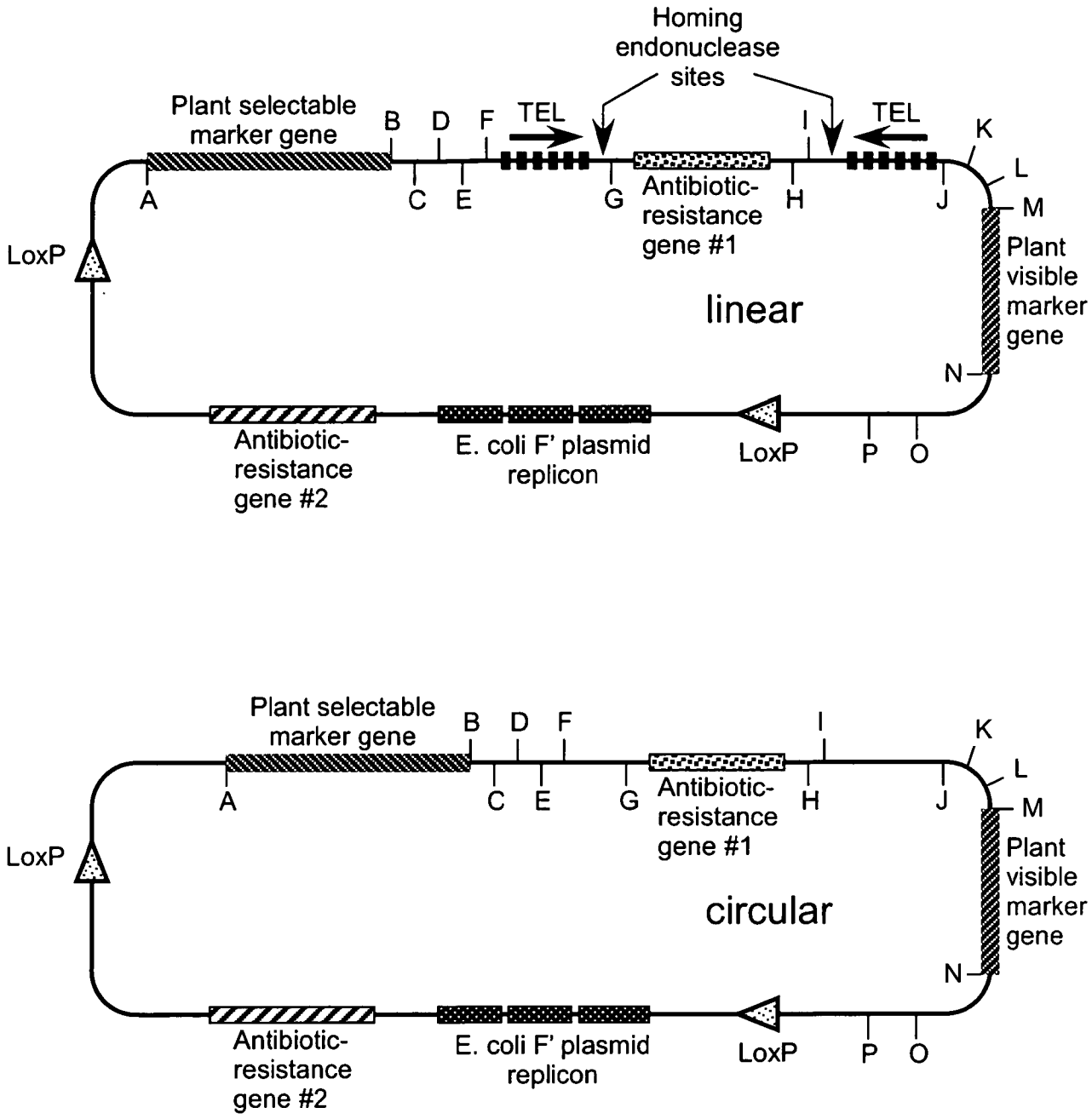
Homing endonuclease sites
Plant selectable marker gene
TEL
TEL
A
B
C
D
E
F
G
H
I
J
K
L
M
N
LoxP
Antibiotic-resistance gene #1
linear
Plant visible marker gene
Antibiotic-resistance gene #2
E. coli F' plasmid replicon
LoxP
P
O
Plant selectable marker gene
circular

Figure 3

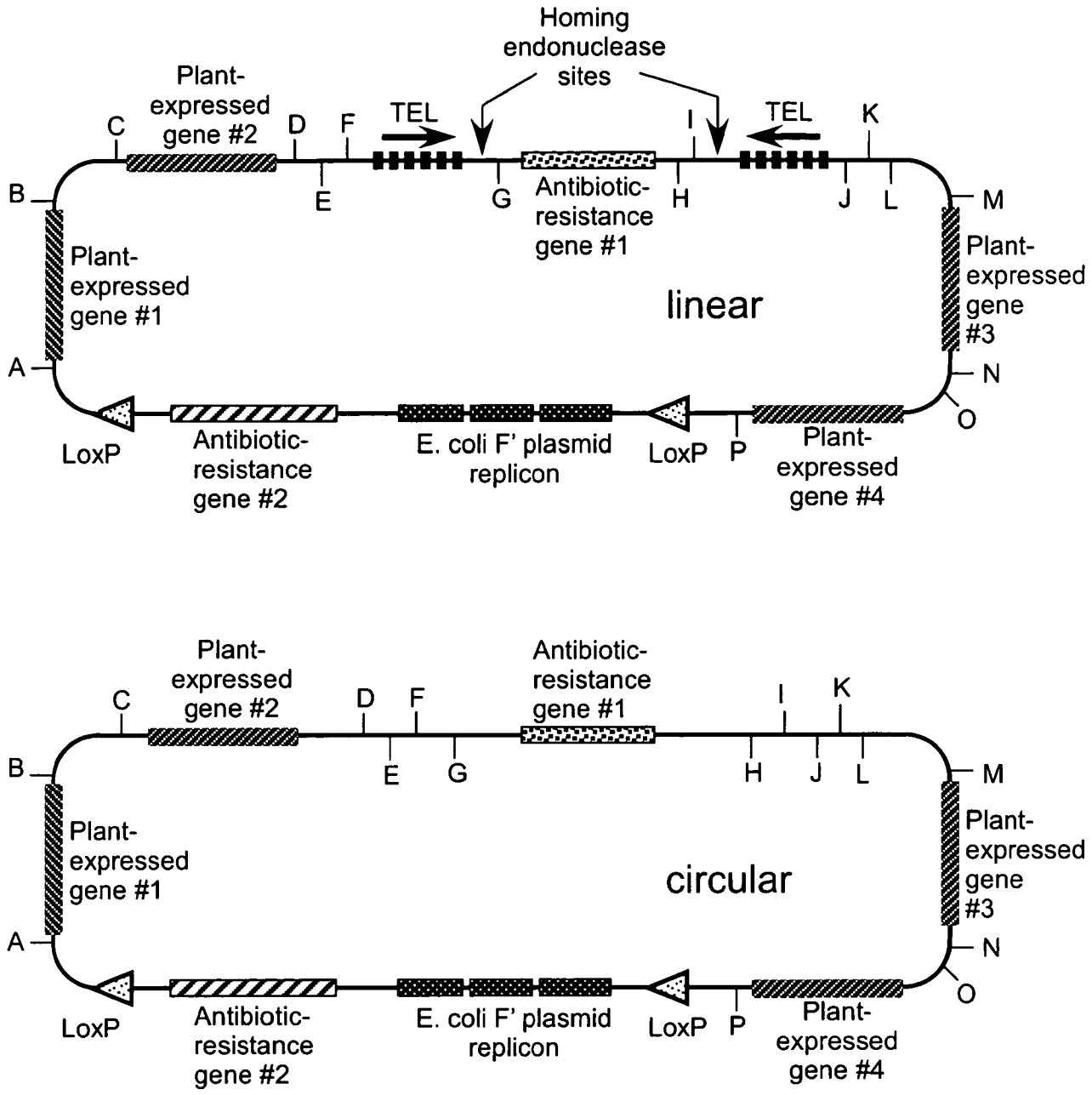
Homing endonuclease sites
Plant-expressed gene #2
TEL
TEL
C
D
F
I
K
B
E
G
Antibiotic-resistance gene #1
H
J
L
M
Plant-expressed gene #1
linear
Plant-expressed gene #3
A
N
O
LoxP
Antibiotic-resistance gene #2
E. coli F' plasmid replicon
LoxP
P
Plant-expressed gene #4
Plant-expressed gene #2
Antibiotic-resistance gene #1
C
D
F
I
K
B
E
G
H
J
L
M
Plant-expressed gene #1
circular
Plant-expressed gene #3
A
N
O
LoxP
Antibiotic-resistance gene #2
E. coli F' plasmid replicon
LoxP
P
Plant-expressed gene #4

Figure 4

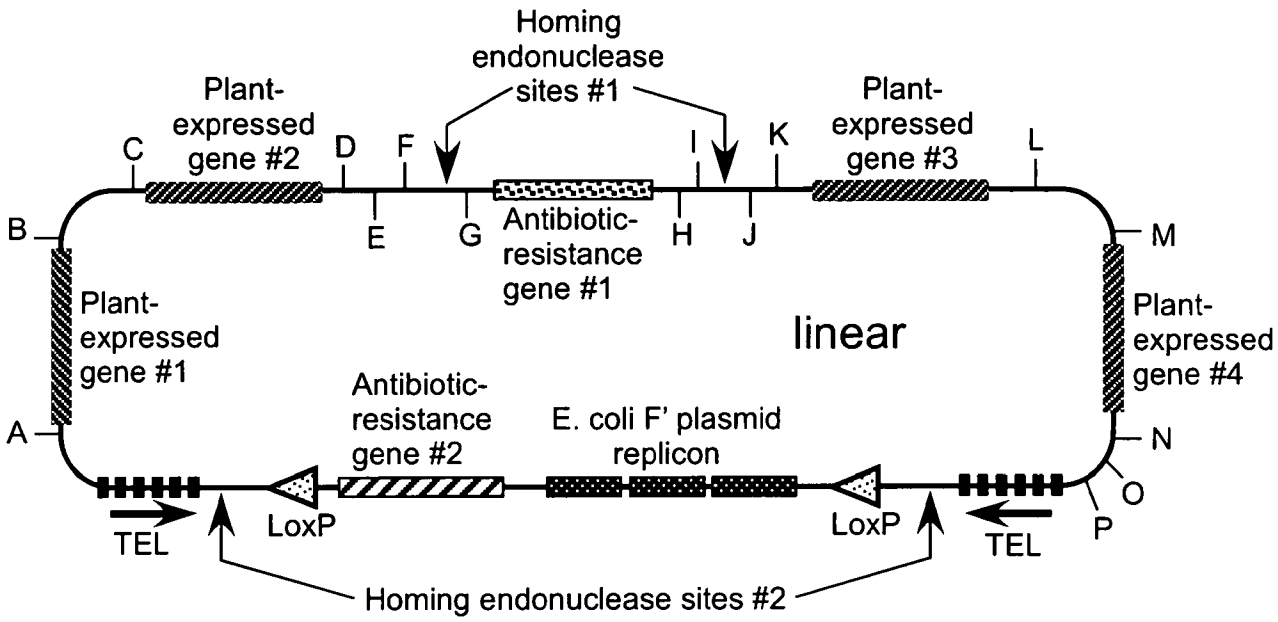
Homing endonuclease sites #1
Plant-expressed gene #2
Plant-expressed gene #3
C
D
F
I
K
L
B
E
G
Antibiotic-resistance gene #1
H
J
M
Plant-expressed gene #1
linear
Plant-expressed gene #4
Antibiotic-resistance gene #2
E. coli F' plasmid replicon
A
N
O
P
TEL
LoxP
LoxP
TEL
Homing endonuclease sites #2

Figure 5

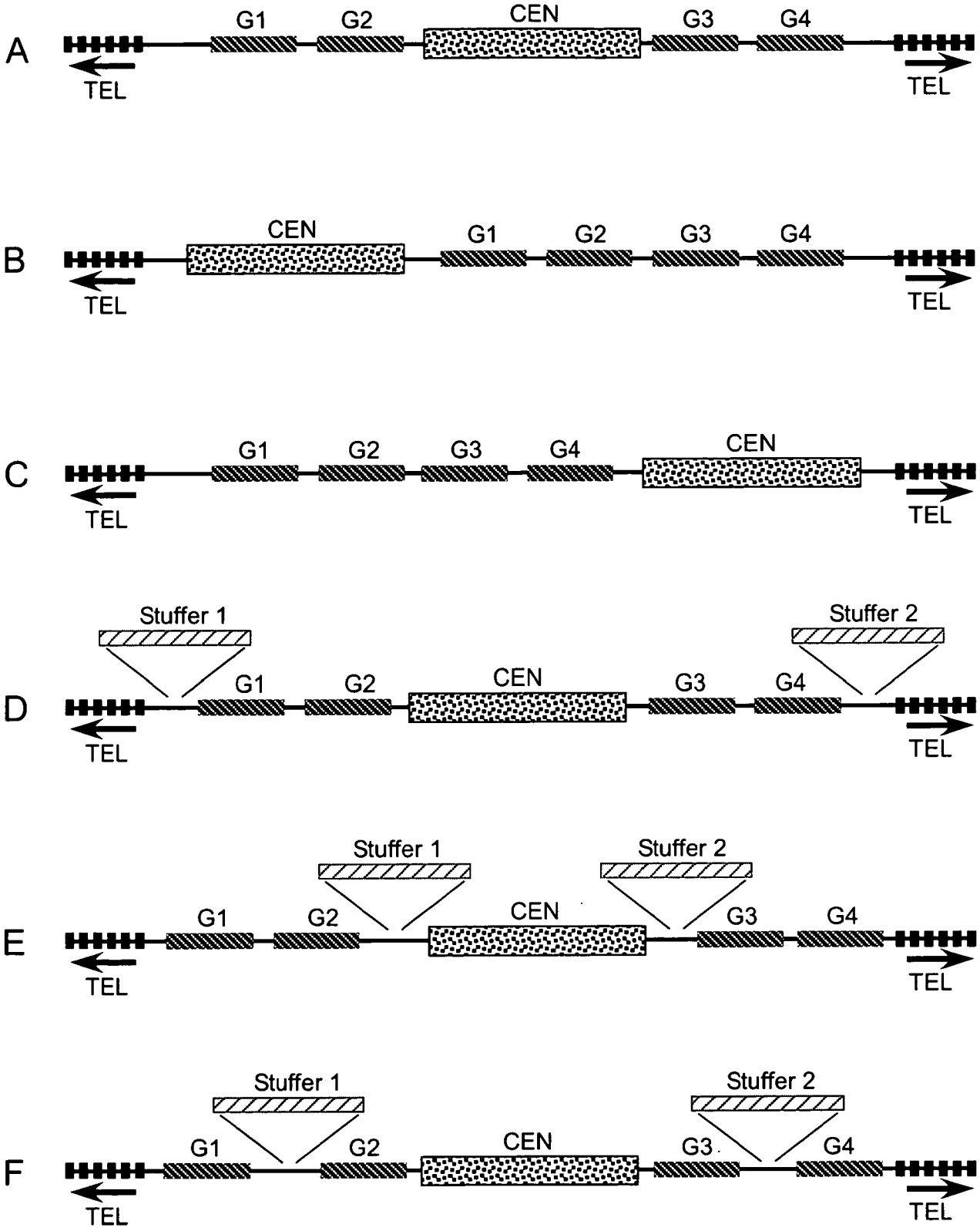
A
G1
G2
CEN
G3
G4
TEL
TEL
B
CEN
G1
G2
G3
G4
TEL
TEL
C
G1
G2
G3
G4
CEN
TEL
TEL
D
Stuffer 1
Stuffer 2
G1
G2
CEN
G3
G4
TEL
TEL
E
Stuffer 1
Stuffer 2
G1
G2
CEN
G3
G4
TEL
TEL
F
Stuffer 1
Stuffer 2
G1
G2
CEN
G3
G4
TEL
TEL

Figure 6

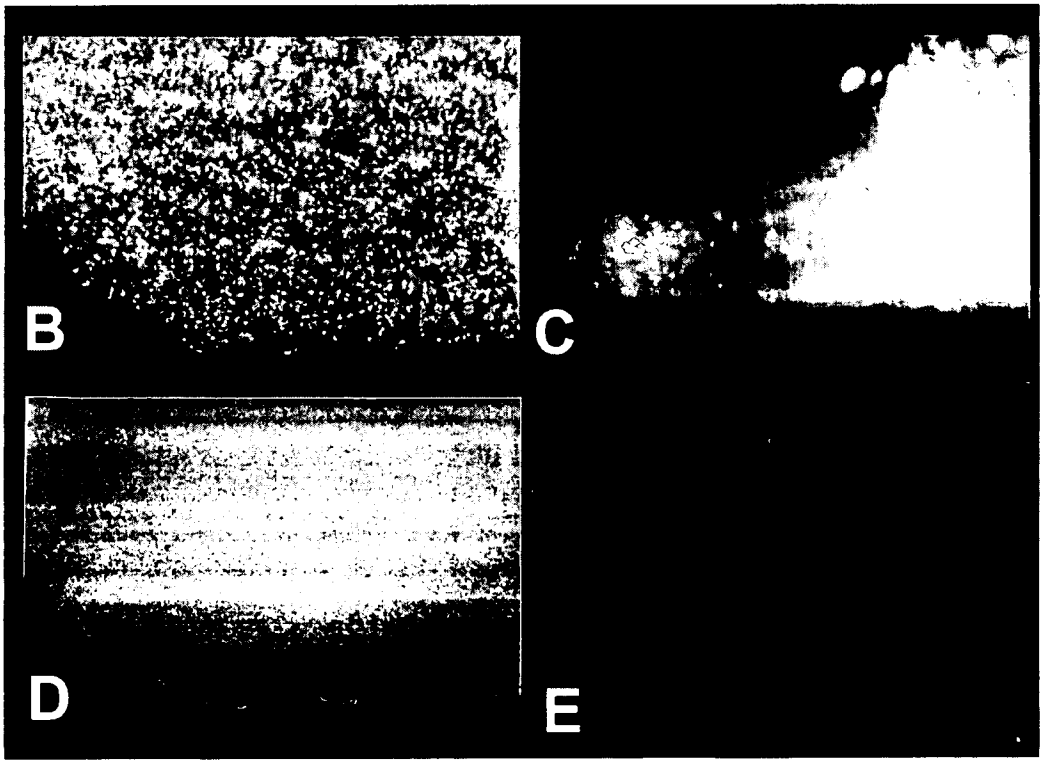

Figure 7